

UNIVERSIDAD COMPLUTENSE DE MADRID

FACULTAD DE CIENCIAS QUÍMICAS



TESIS DOCTORAL

Producción de progesterona por fermentación de esteroides naturales
mediante ingeniería metabólica de sistemas

MEMORIA PARA OPTAR AL GRADO DE DOCTOR

PRESENTADA POR

Gabriel Hernández Fernández

DIRIGIDA POR

José Luis García López
Beatriz Galán Sicilia

UNIVERSIDAD COMPLUTENSE DE MADRID
FACULTAD DE CIENCIAS QUÍMICAS

Programa de doctorado:
Bioquímica, Biología Molecular y Biomedicina



TESIS DOCTORAL

Producción de progesterona por fermentación de esteroides naturales
mediante ingeniería metabólica de sistemas

MEMORIA PARA OPTAR AL GRADO DE DOCTOR PRESENTADA POR

Gabriel Hernández Fernández

DIRECTORES

Prof. José Luis García López

Dra. Beatriz Galán Sicilia

CENTRO DE INVESTIGACIONES BIOLÓGICAS MARGARITA SALAS

CONSEJO SUPERIOR DE INVESTIGACIONES CIENTÍFICAS



Madrid, 2025

El trabajo descrito en esta Tesis Doctoral se ha llevado a cabo en el Grupo de Biotecnología Medioambiental, perteneciente al Departamento de Biotecnología del Centro de Investigaciones Biológicas “Margarita Salas” del Consejo Superior de Investigaciones Científicas (CIB-CSIC) (Madrid, España).

La investigación ha sido financiada por el contrato de Formación de Profesorado Universitario del Plan Estatal de Investigación Científica y Técnica y de Innovación 2021-2023 del Ministerio de Universidades (FPU21/05101) y por el proyecto FRONTEST: Retos en la frontera biotecnológica para la producción de esteroides (PID2021-125370OB-I00) del Ministerio de Ciencia e Innovación.

AGRADECIMIENTOS

Desde que empecé esta Tesis Doctoral he ido pensando en cómo quedaría el que será el apartado más leído y, aun así, es tanto lo que tengo que agradecer a cada uno de vosotros que me cuesta elegir las mejores palabras que os merecéis.

En primer lugar, muchas gracias a Bea por ser la primera en confiar en mí. Gracias a ti me sacaste un poquito de mi zona de confort y, no solo me animaste así a empezar el doctorado, sino también me has empujado a sacar lo mejor de mí cuando yo no lo veía posible. Has estado siempre conmigo al pie del cañón con cada experimento, aportando ideas y soluciones cuando las necesitaba y, al mismo tiempo, manteniendo tu actitud positiva en los experimentos más complicados. Gracias por tu cercanía y tu practicidad cuando me he sentido un poco perdido a lo largo de este trabajo, que no han sido pocas veces. A José Luis le agradezco el invitarme a formar parte de los *Rubenes* cuando en 2018 entré por primera vez al 344. Ha sido una suerte y un orgullo ser de tus últimos doctorandos antes de tu jubilación. Me has enseñado a integrar todos los conocimientos de Bioquímica, Biología molecular y Microbiología con los que uno sale de la carrera y a aplicarlos en el laboratorio. Muchas gracias por dejarme pensar, equivocarme y aprender, por enseñarme a añadir ese punto crítico a todo lo que sabemos y por aportar siempre seguridad y sabiduría cuando nadie más podía hacerlo.

Si yo me lancé al vacío a hacer un doctorado fue simplemente porque sabía que al menos abajo me esperaba el mejor grupo de personas que se puede desear como compañeros de trabajo. Muchas gracias a Paula, JD, Marco, Lara, Isa P., Elena A., Unai, Alberto, Gon, Helena, MSG, Eugenio, Paz, Pablo, Susana, Bea... en definitiva, a todos los *Rubenes* por compartir tanto tiempo conmigo, ya fuera entre pipetas o con una cerveza en la mano. Gracias por dedicar vuestro tiempo a mí y a cualquier persona que necesitaba ayuda, por cualquier risa tonta por los pasillos, en el ascensor o en las comidas y por ser tan buen ejemplo de compañerismo. Gracias además a Vir por echarme una mano con el HPLC y el GC-MS y resolver todas mis curiosidades y a Melissa, Ana V. y Elena R. por hacernos la vida en el laboratorio tremendamente cómoda. A Celia, Andrés, Julia y Víctor, además, os digo que es reconfortante ver que el espíritu de los *Rubenes* sigue en vosotros y sé que vais a dejar huella en la planta 3. También quiero agradecer a los jefes de los *Rubenes* Auxi, Eduardo y Jesús por tener siempre su puerta abierta para resolver cualquier duda o para rebuscar entre glicerolos antiguos, por aportar otros puntos de vista a nuestros seminarios y también por tener tan buen ojo a la hora de contratar y ampliar el grupo que hemos formado.

Específicamente hay una serie de personas por las que, sin ellas, esto no hubiera sido lo mismo. Muchas gracias a Kris por ser mi alma gemela. Todos tus abrazos, tus sonrisas y el tiempo compartido contigo han sido un regalo. Siempre recordaré llegar cada mañana al CIB y verte en tu sitio sin parar de trabajar, pero dispuesta a escucharme. Eres de las personas con el corazón más bonito que conozco y te agradezco mucho todo lo que me quieres (y ya sabes que es recíproco). El tiempo que estuviste en el 345 eras un refugio y, a pesar de que dejaras tu tesis y los *Rubenes*, siempre te he sentido a mi lado, así que parte de esta tesis también es tuya.

A Sofi por ser uno de los motivos por los que volvía con ganas cada día al trabajo. No me creo la suerte que tengo de haberte conocido y haber compartido todo este tiempo los dos. Gracias por tu sonrisa, por estar siempre atenta de mí, por todas las cervezas, los viajes, las comidas y las bromas juntos. Gracias por formar conmigo el Eje del mal y vivir conmigo las cosas más graciosas y, a veces también, las más tristes que me han ido pasando. Creo que pocas personas tan buenas como tú me quedan por conocer en la vida.

Muchas gracias a Santi por ser la otra piedra angular del Eje del mal. Fuera del laboratorio has marcado todos los viernes de cervezas, los viajes y las liadas juntos, mientras que dentro del CIB siempre apreciaré que has estado ahí para escucharme, hacerme reír e intentar que dejara de estar tristón. Tienes un corazón muy noble y acogedor y admiro mucho todo el interés y las ganas de saber que pones en la investigación.

A Sergio por ser la persona con más energía positiva que conozco. Estar contigo es una alegría, tienes una personalidad única que nos ha enamorado a todos y además sabes que me iría contigo al fin del mundo. Muchas gracias por tener siempre una sonrisa para mí y ser uno de mis apoyos en todo momento.

Junto a Sofi, Santi y Sergio, tengo que dar también las gracias a Carla, Alba, Marina y Carlos por hacerme olvidar durante un ratito lo duro que se hace el trabajo de laboratorio mientras comemos patatas y tiramos dados. Todas las tardes pasadas con vosotros han sido fantásticas.

Muchas gracias a Juan por enseñarme todo lo que sé sobre *C. tardaugens*, por acogerme en tu despacho para consolarme en mis fracasos y salir de ahí con la motivación renovada y una sonrisa. Te honra la buena voluntad que tienes para cambiar las cosas a mejor y allanar el camino a los que vienen detrás. Te admiro mucho y eres un ejemplo a seguir.

En cuanto a toda la gente que ha pasado por el 344 y ha compartido conmigo no solo espacio, sino también temperaturas extremas, muchas gracias por vivir la parte más experimental de mi tesis. Muchas gracias a Irene por ser mi compañera de espalda todos estos años. Por compartir el humor más tonto cuando ya estábamos agotados mentalmente, por decorar nuestro pequeño rincón y por hacer excursiones juntos a comer todos los dulces que hay disponibles, no hubiera elegido otra compañera de tesis. A María por ser mi primera mamá pata, recibirme con cariño y por toda la música escuchada juntos y a Sara por llenar el 344 de risas y anécdotas. Las tres sois mi miércoles lila. Muchas gracias a Miguel por haber compartido conmigo la única neurona que nos quedaba. Por todas esas mañanas hablando de nuestras tonterías, con ese humor que solo puedo tener contigo, y por todos tus consejos. A Isa por continuar los buenos ratos en el laboratorio y hacerme más amenos los ratos de espera de los experimentos, siempre dispuesta a sacarme un tema de conversación. A los dos os tengo que dar las gracias por toda la ayuda con los reactores y los consejos para los artículos. Me habéis animado un montón para conseguir llegar hasta aquí. Finalmente, muchas gracias a Juanjo por haber traído al

laboratorio muchas risas y muchas chuches y perdón por no apuntar bien y tirar en tu poyata todas las puntas.

Dejando ya a los *Rubenes*, pero siguiendo en el CIB, muchas gracias a los *Funquitos* (Carlos, Ale, Diego, Javi, Felipe, Ana...) por ser un grupo tan divertido y compartir cervezas, Fiestas de la Primavera y casas rurales con nosotros.

Dejando atrás el mundo de la ciencia, ahora me toca agradecer a los que lleváis conmigo desde un principio y me habéis formado y hecho crecer como persona. A mis padres por cuidarme y educarme, enseñarme lo que es la bondad y la empatía, darme valores y apoyarme en todas mis decisiones. Vosotros habéis puesto las primeras piedras sobre las que se ha construido todo lo que soy. A mi hermana por estar siempre a mi lado, ser mi confidente y protegerme. No hubiera podido enfrentarme a muchas cosas sin saber que al volver a casa estarías tú si algo salía mal. A mis amigos del barrio por haber crecido juntos y vivir miles de anécdotas. Muchas gracias por haber visto todas las versiones de mí y decidir seguir conmigo. Sois el núcleo de mi corazón y la fuente de todas mis risas. A mis amigos de la uni por acogerme en uno de los momentos más vulnerables de mi vida y compartir conmigo el interés por la Bioquímica. Y, finalmente, a Nacho y a Álex. Muchas gracias a ambos por dejarme conocerlos y hacerme un huequito en vuestros corazones, vosotros os habéis ganado el mío por completo. Me habéis escuchado, me habéis cuidado y, sobre todo, me habéis entendido. Como dijo Álex, vuestras risas armonizan, y yo solo quiero pasar las tardes riendo con vosotros. Sois los mejores.

Esta tesis tiene un poco de todos vosotros.

Índice

Abreviaturas	1
I. Resumen.....	5
II. Summary.....	11
III. Introducción	17
1. Introducción a los esteroides	19
1.1. Esteroles	19
1.2. Hormonas esteroideas	20
1.3. Los esteroides en la industria farmacéutica.....	22
2. Biotecnología de los esteroides	22
2.1. Metabolismo microbiano de los esteroides	24
2.1.1. Transporte, catabolismo y regulación en actinobacterias	24
2.1.2. Metabolismo del colesterol en bacterias Gram negativas.....	28
2.2. Producción microbiana de sintonas esteroideas	30
3. Los esteroides como contaminante medioambiental.....	33
3.1. Degradación de esteroides con actividad EDC	35
3.1.1. Degradación aeróbica de andrógenos	36
3.1.2. Degradación aeróbica de progestágenos.....	37
3.1.3. Degradación aeróbica de estrógenos.....	39
4. El futuro de la biotecnología de esteroides.....	40
IV. Objetivos	43
V. Materiales y métodos	47
1. Cepas bacterianas, plásmidos, oligonucleótidos y reactivos	49
2. Medios y condiciones de cultivo	63
2.1. Medios y condiciones de cultivo para <i>Escherichia coli</i>	63
2.2. Medios y condiciones de cultivo para <i>Caenibius tardaugens</i>	63
2.3. Medios y condiciones de cultivo para <i>Mycolicibacterium smegmatis</i>	63
2.4. Preparación de esteroides	64
2.5. Conservación de las cepas bacterianas.....	64
3. Técnicas de manipulación de ADN	64
3.1. Extracción de ADN cromosómico	65
3.2. Aislamiento de ADN plasmídico	65

3.3.	Reacción en cadena de la polimerasa de ADN (PCR)	65
3.4.	Electroforesis de ADN en gel de agarosa.....	66
3.5.	Secuenciación de ADN	66
3.6.	Clonaje <i>Golden Standard</i>	66
4.	Experimentos de modificación genética.....	67
4.1.	Transformación de células de <i>E. coli</i>	67
4.2.	Transformación de células de <i>C. tarDAUGENS</i>	67
4.3.	Transformación de células de <i>M. smegmatis</i>	68
4.4.	Obtención de mutantes por delección génica.....	68
5.	Técnicas de manipulación de ARN	69
5.1.	Extracción de ARN total de <i>C. tarDAUGENS</i>	69
5.2.	Secuenciación masiva de ARN (<i>RNA-seq</i>)	69
6.	Técnicas de manipulación de proteínas	70
6.1.	Obtención de extractos proteicos de <i>E. coli</i>	70
6.2.	Electroforesis en geles de poliacrilamida-SDS (PAGE-SDS)	70
6.3.	Ensayos enzimáticos <i>in vitro</i> de la actividad Baeyer-Villiger monooxigenasa (BVMO)	71
6.4.	Ensayos enzimáticos <i>in vitro</i> de la actividad de citocromos CYP450	71
6.5.	Modelado y análisis <i>in silico</i> de la estructura de los CYP450 candidatos de <i>C. tarDAUGENS</i>	71
7.	Biotransformaciones de esteroides <i>in vivo</i>	72
7.1.	Ensayos de biotransformación de esteroides con células en crecimiento en matraz.	72
7.2.	Ensayos de biotransformación de esteroides con células en reposo (<i>resting cells</i>) en matraz.....	72
7.3.	Ensayos de biotransformación de esteroides en crecimiento a escala de biorreactor	73
8.	Técnicas cromatográficas	73
8.1.	Extracción de esteroides.....	73
8.2.	Cromatografía líquida de alta eficiencia (HPLC) y acoplada a espectrometría de masas (HPLC-MS)	74
8.3.	Cromatografía de gases acoplada a espectrometría de masas (GC-MS).....	74
9.	Ensayos de fluorescencia en placas multipocillo	75
10.	Cálculos.....	75
11.	Análisis bioinformáticos	75

VI. Resultados 77

Capítulo 1. Producción de esteroides con interés industrial en *Mycolicibacterium smegmatis* a partir de esteroides 79

- 1.1. Optimización de la producción de androstenediona (C19) mediante biotransformación a partir de esteroides 81
- 1.2. Producción del esteroide C22 22-hidroxi-23,24-bisnorcol-4-en-3-ona mediante biotransformación de esteroides 82
- 1.3. Desarrollo de nuevas herramientas genéticas en micobacterias: expansión del *Golden Standard*..... 83
 - 1.3.1. Construcción de la colección de plásmidos con *ori* RSF1010..... 85
 - 1.3.2. Validación de la nueva colección de *Golden Standard* en *M. smegmatis*.. 86
 - 1.3.2.1. Caracterización de la colección de promotores 86
 - 1.3.2.2. Caracterización de la colección de RBSs 88
 - 1.3.2.3. Caracterización de la colección de terminadores de la transcripción..... 89
 - 1.3.2.4. Caracterización de los sistemas inducibles 90
- 1.4. Producción bacteriana de progesterona 92
 - 1.4.1. Construcción de una cepa de *M. smegmatis* deficiente en el consumo de esteroides. 93
 - 1.4.2. Biotransformación de esteroides en progesterona en *M. smegmatis* 95
 - 1.4.3. Mejora de la producción de progesterona mediante *Golden Standard*..... 100

Capítulo 2. Estudio del metabolismo de la progesterona en *C. taradagens* NBRC 16725 103

- 2.1. Metabolismo de la progesterona en *C. taradagens* 105

Capítulo 3. Estudio del catabolismo del colesterol en *C. taradagens* NBRC 16725 107

- 3.1. Análisis del crecimiento de *C. taradagens* en colesterol 109
- 3.2. Actividad del citocromo CYP2113 110
- 3.3. Caracterización del mutante Δ LLM-BVMO en colesterol 112
- 3.4. Aplicación de la cepa Δ LLM-BVMO para producir progesterona a partir de colesterol 115
- 3.5. Análisis *in silico* de CYP450, Fdx y FdxR inducidos en colesterol 116
- 3.6. Análisis *in silico* de otros genes inducidos en colesterol 119

VII. Discusión 123

- 1. Utilización de *M. smegmatis* como factoría celular para la producción de esteroides con interés industrial 125

1.1. Producción de esteroides C19 y C22.....	125
1.2. Ampliación de la herramienta genética <i>Golden Standard</i> para micobacterias y su aplicación en la producción de progesterona.....	127
2. Metabolismo de la progesterona en <i>C. tardaugens</i>	133
3. Metabolismo del colesterol en <i>C. tardaugens</i>	137
VIII. Conclusiones	145
IX. Bibliografía	149
X. Anexos	169

Abreviaturas

1,4-HBC:	22-hidroxi-23,24-bisnorcol-1,4-dien-3-ona
16-dPREG-Ac:	acetato de 16-dehidropregnenolona
17 β -HSD:	17 β -hidroxiesteroide deshidrogenasa
20,22-dOHCHO:	20R,22R-dihidroxicolesterol
22-OBNC-CoA:	3,22-dioxo-colest-4-en-24-carboxil-CoA
22-OH-BCN-CoA:	22-hidroxi-3-oxo-colest-4-en-24-carboxil-CoA
22-OHCHO:	22R-hidroxicolesterol
3-HSA:	3-hidroxi-9,10-secoandrosta-1,3,5(10)-trien-9,17-diona
3-OPA:	(20S)-3-oxopregn-4-en-20-carboxaldehido
3-OPC-CoA:	3-oxo-4-pregnen-20-carboxil-CoA
3 β -HSD:	3 β -hidroxi- Δ 5-esteroide deshidrogenasa/isomerasa
4,9-DSHA: oico	4,5,9,10-diseco-3-hidroxi-5,9,17-trioxoandrosta-1(10),2-dien-4-
4-HBC:	22-hidroxi-23,24-bisnorcol-4-en-3-ona
4-OHE1:	4-hidroxiestrone
9-OHAD:	9 α -hidroxi-1,4-androstadien-3,17-diona
9-OHHBC:	9,22-dihidroxi-23,24-bisnorcol-4-en-3-ona
ABC:	<i>ATP binding cassette</i>
AD:	4-androsten-3,17-diona
ADD:	1,4-androstadien-3,17-diona
Adx:	adrenodoxina
AdxR:	adrenodoxina reductasa
aTC:	anhidrotetraciclina
BLAST:	<i>Basic local allingment search tool</i>
BPA:	Bisfenol A
BSTFA:	N,O-bis(trimetilsilil)trifluoroacetamida con trimetilclorosilano
BVMO:	Baeyer-Villiger monooxigenasa
CDO:	colest-1,4-dien-3-ona

CDS:	secuencias codificantes
CDX:	metil- β -ciclodextrina
CH/L:	colesterol hidroxilasa/liasa.
CHO:	colesterol
CHONA:	colesteno
ChOx:	colesterol oxidasa
CYP450sc:	citocromo P450 <i>side chain cleavage</i>
DDO-26-CoA:	derivado CoA del ácido DDO-26-oico
DES:	dietilestilbestrol
DH:	deshidrogenasa
DMSO:	dimetilsulfóxido
dNTP:	desoxinucleótido trifosfato
DO ₆₀₀ :	densidad óptica a 600 nm
DSAO:	1,17-dioxo-2,3-secoandrostan-3-oico
E1:	estrone
E2:	estradiol
E3:	estriol
EDC:	disruptor endocrino
edc:	<i>estrogen degrading cluster</i>
EE2:	etinilestradiol
EEQ:	equivalentes de estradiol
FAD:	flavín adenín dinucleótido
Fdx:	ferredoxina
FdxR:	ferredoxina reductasa
FMN:	flavín mononucleótido
FS:	sitios de fusión
GC-MS:	cromatografía de gases acoplada a espectrometría de masas
GFP:	proteína verde fluorescente
GmR:	resistencia a gentamicina

GS:	<i>Golden Standard</i>
HDD:	2-hidroxihexa-2,4-dienoico
HIP:	dioxo-1,2,3,4,10,19-hexanorandrostan-5-oico
HPLC-MS:	cromatografía líquida de alta eficiencia acoplada a espectrometría de masas
Hsd4A:	β -hidroxiacil-CoA deshidrogenasa
IPTG:	isopropil- β -D-tiogalactopiranosido
kb:	kilobase
kDa:	kiloDalton
KmR:	resistencia a kanamicina
KshAB:	cetoesteroide-9 α -hidroxilasa
KstD:	3-cetoesteroide- Δ^1 -deshidrogenasa
LB:	medio líquido Luria-Bertani
LLM:	<i>luciferase-like monooxygenase</i>
MCE:	<i>mammalian cell entry</i>
MEM:	membrana externa mitocondrial
MIM:	membrana interna mitocondrial
MoClo:	<i>Modular Clonning</i>
MY:	rendimiento molar
NAD(H):	nicotinamida-adenina-dinucleótido
NADP(H):	fosfato de nicotinamida-adenina-dinucleótido
NB:	<i>nutrient broth</i>
p/v:	peso/volumen
PAGE-SDS:	electroforesis en geles de poliacrilamida-SDS
pb:	pares de base
PDB:	<i>Protein Data Bank</i>
pdC:	<i>progesterone degradation cluster</i>
PHB:	poli(3-hidroxibutirato)
PHY:	fitoesteroles
PREG:	pregnenolona

Abreviaturas

PROG:	progesterona
RBS:	sitios de unión al ribosoma
RFP:	proteína roja fluorescente
RIN:	<i>RNA integrity number</i>
SD:	<i>steroid degrading</i>
SDS:	dodecilsulfato sódico
SmR:	resistencia a espectinomicina
ST:	estándar
TE:	Tris-EDTA
TES:	testosterona
TPM:	tránscrios por millón
URF:	Unidades relativas de fluorescencia
USD:	dólar estadounidense
v/v:	volumen/volumen
X-Gal:	5-bromo-4-cloro-3-indolil- β -D-galactopiranósido



I. RESUMEN

Producción de progesterona por fermentación de esteroides naturales mediante ingeniería metabólica de sistemas

Introducción

Los esteroides presentan funciones muy diversas, ya sea desempeñando un papel estructural en membranas celulares como el colesterol, o presentando actividad hormonal como la progesterona o la cortisona, y por ello los fármacos esteroideos se encuentran incluidos en la lista de medicamentos cruciales por la OMS. Específicamente, la progesterona y sus derivados no solo se utilizan para el tratamiento de la endometriosis o como método anticonceptivo, sino que sirven como punto de partida para producir otros fármacos, como el desoxicortisol y la hidrocortisona.

La progesterona se produce actualmente a partir de la diosgenina empleando 8 pasos de síntesis química, resultando en un proceso contaminante y con un rendimiento considerablemente bajo, generando la necesidad de buscar una alternativa que reduzca tanto su coste como su impacto medioambiental. Esta situación abre la posibilidad de utilizar microorganismos que realicen todo o parte del proceso de producción de los esteroides de manera más eficiente y menos contaminante. La semisíntesis de esteroides utilizaría primero una biotransformación bacteriana de un sustrato esteroideo de partida (p.e., esteroides naturales) en una molécula intermediaria, o sintona esteroidea, para luego obtener el fármaco de interés con menos pasos químicos.

Una semisíntesis alternativa ya aplicada a escala industrial es el uso de micobacterias, como *Mycolicibacterium smegmatis* y otras especies, que transforman esteroides naturales en una sintona C22 llamada 22-hidroxi-23,24-bisnorcol-4-en-3-ona (4-HBC) la cual, mediante solo 2 pasos de síntesis química, permite obtener progesterona. Sin embargo, esta producción requiere todavía de reactivos químicos y las biotransformaciones actualmente presentan algunos retos. Un ejemplo es que, al producir sintonas C22, también se acumulan esteroides C19, y viceversa. Esto disminuye el rendimiento del proceso y dificulta las purificaciones posteriores. Por eso aún se requiere o bien mejorar la productividad de las factorías celulares actuales, o bien buscar nuevas rutas de biosíntesis de estos esteroides.

Una alternativa para la producción directa y bio-basada de progesterona es la biotransformación de esteroides utilizando los citocromos CYP450_{scc}. Estos CYP450_{scc} hidroxilan la cadena lateral de los esteroides y la cortan, produciendo progesterona sin la necesidad de emplear reactivos químicos. Además de los CYP450_{scc} animales, se han descrito dos CYP450_{scc} bacterianos provenientes de *Novosphingobium aromaticivorans* y *Novosphingobium subterraneum* que poseen un gran potencial para la producción de progesterona a nivel industrial, pero cuya aplicabilidad no se ha estudiado en profundidad.

El aumento de la producción y consumo de los esteroides con actividad hormonal supone también un gran problema de contaminación ambiental, ya que una vez liberados al medioambiente presentan actividad disruptora endocrina (EDC) alterando el

metabolismo y la reproducción de los animales. Por ello es importante estudiar la degradación bacteriana de los EDCs esteroideos con el objetivo de encontrar un método eficaz para eliminarlos. En los últimos años se han descrito las rutas de degradación bacteriana de algunos EDCs esteroideos como la testosterona o el estrógeno. Sin embargo, la genética y la bioquímica de las rutas implicadas en la degradación de la progesterona no se han descrito.

La bacteria *Caenibius tardaugs* es un modelo interesante para estudiar las rutas de degradación de EDCs, ya que es capaz de degradar distintos compuestos esteroideos como testosterona, estrógeno, progesterona y colesterol, entre otros, y gracias a la disponibilidad de herramientas genéticas, permite estudiar en detalle el catabolismo de esteroideos con cadena lateral, como la progesterona, los ácidos biliares y el colesterol o sin cadena lateral como la testosterona o el estradiol.

Objetivos

El principal objetivo de esta Tesis Doctoral es avanzar en el conocimiento del metabolismo bacteriano de los esteroides para afrontar la producción de nuevas sintonas esteroideas en factorías celulares y conocer las rutas de degradación de EDCs para utilizar los microorganismos degradadores como posibles agentes biorremediadores.

Para cumplir este objetivo, al inicio de esta Tesis Doctoral se plantearon las siguientes objetivos o tareas:

1. Producir progesterona y otras sintonas esteroideas de interés industrial mediante fermentación a partir de esteroides, utilizando *M. smegmatis* como factoría celular y optimizar su producción mediante el uso de la ingeniería genética.
2. Dilucidar la ruta de degradación de la progesterona en *C. tardaugs*.
3. Dilucidar la ruta de degradación del colesterol en *C. tardaugs* y establecer una posible relación con el catabolismo de la progesterona.

Resultados

Objetivo 1. Producción de progesterona y sintonas esteroideas en *M. smegmatis*

La producción de sintonas C19 sin acumulación de subproductos C22 se consiguió mutando el gen *MSMEG_6561*, que codifica la aldolasa msSal, en la cepa MS6039-5941 de *M. smegmatis* productora de androstenediona.

Para mejorar la producción de sintonas C22, se sobreexpresó el gen *MSMEG_6561* (msSal) junto con los genes *MSMEG_6563* y *MSMEG_1623*, que codifican las reductasas msRed y msOpccR respectivamente, en la cepa MS6039-5941-5903 productora de 4-HBC.

Se han desarrollado nuevas herramientas genéticas con el objetivo de controlar de manera más precisa la expresión de los genes homólogos o heterólogos clonados en este hospedador.

Se ha abordado la producción de progesterona bio-basada a partir de esteroides expresando el citocromo CYP450_{scc}, la ferredoxina y la ferredoxina reductasa de *N. aromaticivorans*, capaces de romper la cadena lateral del colesterol y la colestenoína y formar pregnenolona y progesterona, respectivamente, en la cepa Δ Cyp125A3- Δ 142A2- Δ 125A4 de *M. smegmatis* y se han ensayado distintas condiciones para optimizar la biotransformación.

Objetivo 2. Estudio de la ruta de degradación de la progesterona en *C. tardaugens*

El estudio comparativo del transcriptoma completo de *C. tardaugens* cultivado en progesterona frente a testosterona permitió la identificación del *cluster* génico *pdc*, una agrupación de genes con elevados niveles de expresión en presencia de progesterona entre los que se encontraron las monooxigenasas BVMO y LLM codificadas por los genes *EGO55_13845* y *EGO55_13860*, respectivamente. La delección simultánea de ambos genes impidió el crecimiento de *C. tardaugens* en progesterona. Mediante la realización de ensayos enzimáticos se demostró que ambas monooxigenasas catalizan una oxidación de Baeyer-Villiger transformando la progesterona en acetato de testosterona, que ya puede ser metabolizado por la ruta de la testosterona. Además, se comprobó que estas monooxigenasas actúan también sobre otros esteroides con cadena lateral corta como la 1,2-dehidroprogesterona, la pregnenolona y la 21-hidroxiprogesterona.

Objetivo 3. Estudio de la ruta de degradación del colesterol en *C. tardaugens*

C. tardaugens es capaz de utilizar colesterol o colestenoína como única fuente de carbono y energía solamente cuando se suministran disueltos en ciclodextrinas. La mutación del gen *EGO55_10200*, que codifica un CYP450 y se encuentra inducido en presencia de colesterol, impide el crecimiento de la bacteria en colesterol. Además, se ha relacionado el metabolismo de la progesterona y del colesterol ya que la delección de los genes *EGO55_13845* y *EGO55_13860*, involucrados en el catabolismo de la progesterona, impide también el crecimiento de la cepa en presencia de colesterol, pero este mutante es capaz de acumular 1,2-dehidroprogesterona. Todo ello sugiere que el metabolismo del colesterol se realiza en esta bacteria por un nuevo mecanismo que implica la transformación del colesterol en pregnenolona/progesterona mediante un CYP450_{scc}.

Conclusiones

M. smegmatis es una factoría celular muy versátil para producir distintos tipos de sintonas esteroideas. Se han identificado la aldolasa msSal codificada por el gen

MSMEG_6561 y la reductasa msRed codificada por el gen *MSMEG_6563* como responsables de la formación de 4-HBC. Así, se ha conseguido mejorar la producción de androstenediona a partir de esteroides delecionando el gen *MSMEG_6561* que codifica la aldolasa msSal. Por otro lado, para aumentar la producción de 4-HBC a partir de esteroides en *M. smegmatis*, se ha hiperproducido la aldolasa msSal junto con las reductasas msRed o msOpccR, aunque no se evita completamente la acumulación de esteroides C19 como subproducto.

Se ha completado la colección del *Golden Standard* con el *ori* RSF1010 para permitir la aplicación de esta herramienta genética en micobacterias y la ampliación de la librería de promotores, RBS, terminadores y sistemas inducibles.

La expresión heteróloga en *M. smegmatis* del CYP450scc, la ferredoxina y la ferredoxina reductasa de *N. aromaticivorans* permite la producción de progesterona bio-basada en un solo paso a partir de esteroides.

En esta Tesis Doctoral se ha demostrado que *C. tardaugens* es capaz de utilizar la progesterona como única fuente de carbono y energía expresando las monooxigenasas BVMO y LLM codificadas por los genes *EGO55_13845* y *EGO55_13860*, respectivamente. La progesterona es degradada a través del catabolismo de la testosterona al transformarse en acetato de testosterona por una oxidación de Baeyer-Villiger catalizada por ambas enzimas.

En *C. tardaugens* existe una nueva ruta de degradación del colesterol que está conectada con el catabolismo de la progesterona, ya que el doble mutante de los genes *EGO55_13845* y *EGO55_13860* implicados en la degradación de la progesterona no es capaz de crecer en colesterol y acumula el intermediario 1,2-dehidroprogesterona en presencia de este esteroide.



II. SUMMARY

Progesterone production by natural sterols fermentation using system metabolic engineering

Introduction

Steroids present very diverse activities such as playing a structural role in cell membranes like cholesterol or exhibiting hormonal activity like progesterone or cortisone and they are included in the list of crucial drugs determined by the World Health Organization (WHO). Specifically, progesterone and its derivatives are used as a treatment for endometriosis or as a contraceptive and, moreover, progesterone can be used as starting point for the production of other steroidal drugs, such as deoxycortisol and hydrocortisone.

Progesterone is currently produced from the natural sterol called diosgenine through 8 chemical steps, which results in a highly pollutant process with a considerably low yield, generating the need to look for an alternative that reduces both its cost and environmental impact. This situation opens up the possibility of using microorganisms that perform all or part of the steroid production process in a more efficient and less polluting way. Steroid semi-synthesis would first consist on a bacterial biotransformation of the starting steroidal substrate (e.g., natural sterols) into an intermediate molecule, or steroid synthon, and then obtain the drug of interest with fewer chemical steps.

An alternative semi-synthesis already applied on an industrial scale is the use of mycobacteria, such as *Mycolicibacterium smegmatis* and other species, which transform natural sterols into a C22 synthon called 22-hydroxy-23,24-bisnorcholesterol-4-en-3-one (4-HBC) that, through only 2 chemical synthesis steps, allows us to obtain progesterone. However, this production still requires chemical reagents and biotransformations currently present some challenges. One example is that, when producing C22 synthons, C19 steroids are also accumulated, and vice versa. This decreases the yield of the process and makes subsequent purifications more difficult. Therefore, it is still necessary to either improve the productivity of the current cell factories or to look for new biosynthesis routes for these steroids.

An alternative for a direct and bio-based production of progesterone is the biotransformation of sterols using CYP450scc cytochromes. These CYP450scc hydroxylate the sterol side chain and cleave it, producing progesterone without the need for chemical reagents. In addition to animal CYP450scc, two bacterial CYP450scc from *Novosphingobium aromaticivorans* and *Novosphingobium subterraneum* have been described with a great potential for the production of progesterone at the industrial level, but their applicability has not been studied in depth.

The increase in the production and consumption of steroids with hormonal activity is also a major environmental pollution problem, since, once released into the environment, they present endocrine disrupting activity (EDC), altering the metabolism and reproduction of animals. It is therefore important to study the bacterial degradation

of steroidal EDCs in order to find an effective method to eliminate them. In recent years, bacterial degradation pathways of some steroidal EDCs such as testosterone or estrogen have been described. However, the genetics and biochemistry involved in the bacterial progesterone degradation pathway have not been described.

The bacterium *Caenibius tardaugens* is an interesting model to study the catabolism of EDCs, since it is able to degrade different steroid compounds such as testosterone, estrogen, progesterone and cholesterol, among others and, thanks to the availability of genetic tools functional in this strain, the metabolism of steroids with side chain such as progesterone, bile acids and cholesterol or without side chain, such as testosterone or estradiol can be studied in detail.

Objectives

The main objective of this Doctoral Thesis is to advance in the knowledge of the bacterial metabolism of steroids in order to face the production of new steroid synthons in cell factories and to know the EDCs degradation pathways to use the degrading microorganisms as possible bioremediation agents.

Therefore, the following objectives or tasks were proposed at the beginning of this Doctoral Thesis:

1. To produce progesterone and other steroid synthons of industrial interest by sterols fermentation using *M. smegmatis* as a cell factory and to optimize their production through the use of genetic engineering.
2. To elucidate the progesterone degradation pathway in *C. tardaugens*.
3. To elucidate the cholesterol degradation pathway in *C. tardaugens* and establish a possible relationship with progesterone catabolism.

Results

Objective 1. Production of progesterone and steroid synthons in *M. smegmatis*

Production of C19 synthons without accumulation of C22 byproducts was achieved by mutating the *MSMEG_6561* gene, encoding the msSal aldolase, in the AD-producing strain of *M. smegmatis* MS6039-5941.

To enhance the production of C22 synthons, the *MSMEG_6561* (msSal) gene was overexpressed together with the *MSMEG_6563* and *MSMEG_1623* genes, encoding the reductases msRed and msOpccR, respectively, in the 4-HBC-producing strain MS6039-5941-5903.

New genetic tools have been developed with the aim of a more precisely control of the expression of homologous or heterologous genes cloned in this host.

We addressed the production of bio-based progesterone from sterols by expressing the cytochrome CYP450scc, the ferredoxin and the ferredoxin reductase from *N. aromaticivorans*, capable of cleaving the side chain of cholesterol and cholestenone and forming pregnenolone and progesterone, respectively, in the $\Delta Cyp125A3\text{-}\Delta 142A2\text{-}\Delta 125A4$ strain of *M. smegmatis* and different conditions were tested to optimize the biotransformation.

Objective 2. Study of the progesterone degradation pathway in *C. tardaugs*

The comparative study of the complete transcriptome of *C. tardaugs* grown on progesterone versus testosterone allowed the identification of the *pdC* gene cluster, a group of genes with elevated expression levels in the presence of progesterone among which the BMO and LLM monooxygenases are encoded by the *EGO55_13845* and *EGO55_13860* genes, respectively. Simultaneous deletion of both genes prevented the growth of *C. tardaugs* on progesterone. By performing enzymatic assays, it was shown that both monooxygenases catalyze a Baeyer-Villiger oxidation, transforming progesterone into testosterone acetate which can now be metabolized by the testosterone pathway. In addition, these monooxygenases were also able to transform other short side-chain steroids such as 1,2-dehydroprogesterone, pregnenolone and 21-hydroxyprogesterone.

Objective 3. Study of the cholesterol degradation pathway in *C. tardaugs*

C. tardaugs is able to use cholesterol or cholestenone as carbon and energy source only when supplied dissolved in cyclodextrins. Mutation of the *EGO55_10200* gene, which encodes a CYP450 and is induced in the presence of cholesterol, prevents the growth of the bacterium on cholesterol. In addition, progesterone and cholesterol metabolism were found to be linked since the deletion of the *EGO55_13845* and *EGO55_13860* genes, involved in the progesterone catabolism, also prevents the growth of the strain in the presence of cholesterol, and this mutant is able to accumulate 1,2-dehydroprogesterone. This suggests that cholesterol metabolism is carried out in this bacterium by a new mechanism involving the transformation of cholesterol into pregnenolone/progesterone by a CYP450scc.

Conclusions

M. smegmatis is a very versatile cell factory for producing different types of steroid synthons. The aldolase msSal encoded by the *MSMEG_6561* gene and the reductase msRed encoded by the *MSMEG_6563* gene have been identified as responsible for the formation of 4-HBC. Thus, the production of androstenedione from sterols has been improved by deleting the *MSMEG_6561* gene encoding msSal. On the other hand, to increase the production of 4-HBC from sterols in *M. smegmatis*, msSal has been

hyperproduced together with the reductases msRed or msOpccR, although the accumulation of C19 steroids as by-product is not completely prevented.

The Golden Standard collection has been completed with the *ori* RSF1010 that allows the application of this genetic tool in mycobacteria and the expansion of the library of promoters, RBSs, terminators and inducible systems.

Heterologous expression in *M. smegmatis* of the CYP450sc, the ferredoxin and the ferredoxin reductase from *N. aromaticivorans* allows the production of bio-based progesterone in a single step from sterols.

In this Doctoral Thesis it has been demonstrated that *C. tardaogens* is able to utilize progesterone as a sole source of carbon and energy by expressing the monooxygenases BVMO and LLM, encoded by the genes *EGO55_13845* and *EGO55_13860*, respectively. Progesterone is degraded through the testosterone catabolism since it is transformed into testosterone acetate by a Baeyer-Villiger oxidation catalyzed by both enzymes.

A novel cholesterol degradation pathway that is connected to the progesterone catabolism occurs in *C. tardaogens*, since the double mutant strain of *EGO55_13845* and *EGO55_13860* genes involved in progesterone degradation is not able to grow in cholesterol and accumulates the 1,2-dehydroprogesterone intermediate in the presence of this sterol.



III. INTRODUCCIÓN

1. Introducción a los esteroides

Los esteroides son un grupo muy diverso de compuestos orgánicos con múltiples funciones biológicas y con interés medioambiental e industrial. Los esteroides son lípidos triterpenoides tetracíclicos con una estructura central común de 17 átomos de carbono conocida como ciclopentanoperhidrofenantreno, dividida en 3 ciclohexanos llamados anillos A, B y C, y un ciclopentano denominado anillo D (Fig. I1). En función de los distintos grupos funcionales que presenten (grupos metilo, hidroxilo, cetona o cadenas alifáticas) y la posición de estos en la molécula (en C3, C11 o C17, entre otros), estos compuestos poseen diferentes funciones biológicas (Nelson et al., 2013). En esta familia de compuestos podemos encontrar los esteroides, los ácidos biliares, las hormonas esteroideas, derivados de la vitamina D o las saponinas, de las que a su vez derivan las sapogeninas (Olivera y Luengo, 2019).

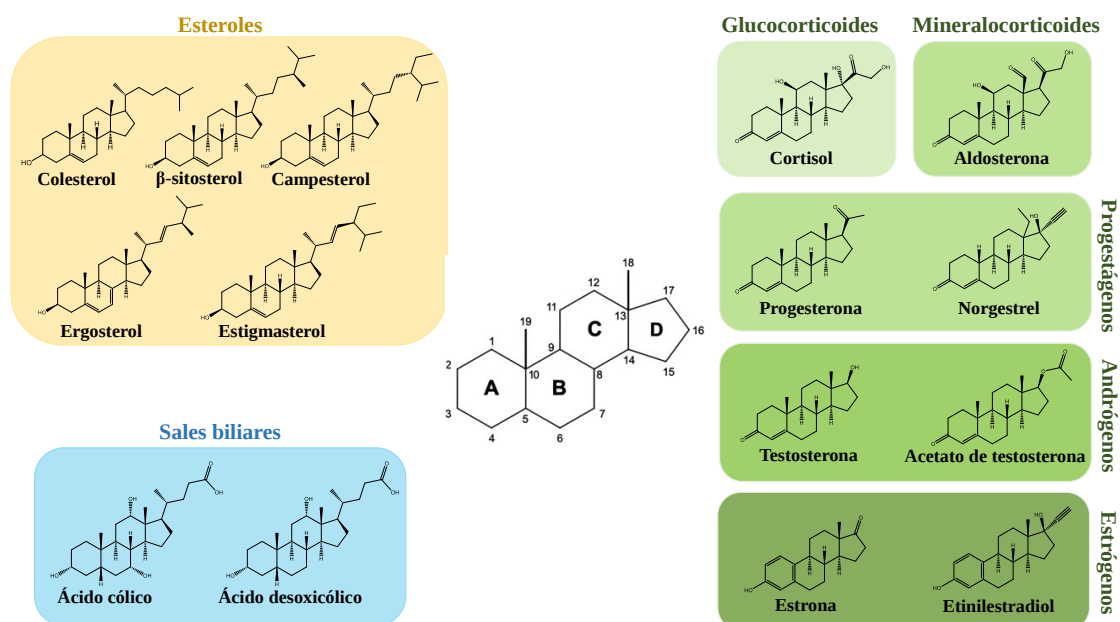


Figura I1. Estructura química de distintos esteroides.

1.1. Esteroides

Los esteroides son un tipo de esteroides caracterizado por presentar un grupo hidroxilo en C3 y una cadena lateral de al menos 8 carbonos en C17. Estos compuestos se encuentran muy distribuidos en la naturaleza entre distintos organismos y entre ellos podemos destacar el colesterol de las células animales, el ergosterol de levaduras y hongos filamentosos y los fitoesteroides de las células vegetales, y cada uno se diferencia según el tipo de cadena alifática en C17 y la posición de los dobles enlaces (Fig. I1) (Parsi y Górecki, 2006; Piironen et al., 2000). La función general de estos esteroides consiste en conferir estabilidad e impermeabilidad a las células eucariotas, siendo por tanto un componente esencial de esta estructura celular (Nelson et al., 2013). A pesar de que la

biosíntesis de estos compuestos se ha considerado un sello de identidad de los organismos eucariotas, se conocen varios casos de síntesis de esteroides (modificados o no) en bacterias metanotróficas como *Methylococcus capsulatus* (Bird et al., 1971; Lamb et al., 2007), en diversas mixobacterias (García et al., 2014; Gawas et al., 2011; Kohl et al., 1983) o en *Gemmata obscuriglobus* (Pearson et al., 2003).

El colesterol es un esteroide con 27 átomos de carbono que no solo actúa como agente estabilizante de membranas, sino que es el precursor de las hormonas esteroideas, ácidos biliares y de la vitamina D en animales (Björkhem y Eggertsen, 2001; Miller, 1988). Este esteroide se puede ingerir a través de la dieta, pero también se sintetiza *de novo* al condensar acetil-CoA en el triterpeneo escualeno, que a su vez se cicla dando lanosterol (Nelson et al., 2013). Los mamíferos no presentan las enzimas necesarias para degradar la estructura central del colesterol de modo que un exceso de este esteroide se excreta a través de la bilis (Björkhem y Eggertsen, 2001). Por esta razón el colesterol o su derivado, el coprostanol, se pueden utilizar como un indicador de contaminación antropogénica en aguas (Frena et al., 2016; Froehner y Sáñez, 2013; Harrauld et al., 2019).

Los esteroides de plantas son los fitoesteroides y los representantes más comunes son el β -sitosterol, el estigmasterol y el campesterol. Los tres se diferencian del colesterol por presentar un grupo etilo o metilo en C24 o un doble enlace entre C22-C23 (Fig. I1) (Piironen et al., 2000). Estos esteroides también derivan del escualeno y sirven además como punto de partida para la síntesis de glicoalcaloides, cardenolidos y saponinas (Piironen et al., 2000).

1.2. Hormonas esteroideas

Las hormonas esteroideas de mamíferos se pueden dividir según el órgano en el que se produzcan. En la corteza de la glándula adrenal se producen los mineralocorticoides como la aldosterona, encargados de controlar la reabsorción de iones inorgánicos en el riñón, y los glucocorticoides como el cortisol (Fig. I1), involucrados en la regulación de la gluconeogénesis y en el control de la respuesta inmune. Por otro lado, las hormonas sexuales se producen en las gónadas masculinas y femeninas y en la placenta, dividiéndose en andrógenos como la testosterona, estrógenos como el estradiol y progestágenos como la progesterona (Fig. I1). Todos ellos están involucrados en la regulación del ciclo reproductivo, el desarrollo de los caracteres sexuales secundarios y en la formación de los gametos (Nelson et al., 2013). Aparte de estas hormonas sexuales de origen natural, en clínica se utilizan derivados sintéticos de andrógenos, estrógenos y progestágenos con actividad endocrina mejorada, como el acetato de testosterona, el cipionato de testosterona, el etinilestradiol o el norgestrel, entre otros (Fig. I1) (Cagnacci et al., 2025).

Todas las hormonas esteroideas son efectivas a muy baja concentración y por eso se sintetizan en bajas cantidades en el organismo. Estos esteroides se forman a partir del colesterol en un proceso denominado esteroidogénesis en el que la cadena lateral se elimina de forma total o parcial y el núcleo esteroideo sufre hidroxilaciones o

deshidrogenaciones (Nelson et al., 2013; Toor y Sikka, 2017). El citocromo CYP11A1 es un CYP450scc (del inglés *side chain cleavage*) que se encuentra en la membrana mitocondrial de mamíferos y actúa junto a la adrenodoxina (Adx) y la adrenodoxina reductasa (AdxR) formando el sistema colesterol hidroxilasa/liasa (CH/L). Este sistema primero hidroxila el C22 del colesterol, originando el 22R-hidroxicolesterol (22-OHCHO), y después el C20, obteniendo así el 20R,22R-dihidroxicolesterol (20,22-dOHCHO) (Slominski et al., 2015; Strushkevich et al., 2011). Esta molécula sufre a continuación una rotura oxidativa generando la pregnenolona y el aldehído isocaproal (Fig. I2). Una vez sintetizada, la pregnenolona es el precursor del resto de hormonas esteroideas, pudiendo ser oxidada y formar progestágenos, hidroxilada para generar los mineralocorticoides y los glucocorticoides o puede perder su cadena lateral para formar andrógenos y estrógenos (Toor y Sikka, 2017). No se ha descrito todavía la degradación completa de estas hormonas en mamíferos, de forma que sufren algunas modificaciones estructurales, como puede ser la conjugación con ácido glucurónico o con sulfato, para ser excretadas por las heces a través de la bilis o por la orina (Nelson et al., 2013).

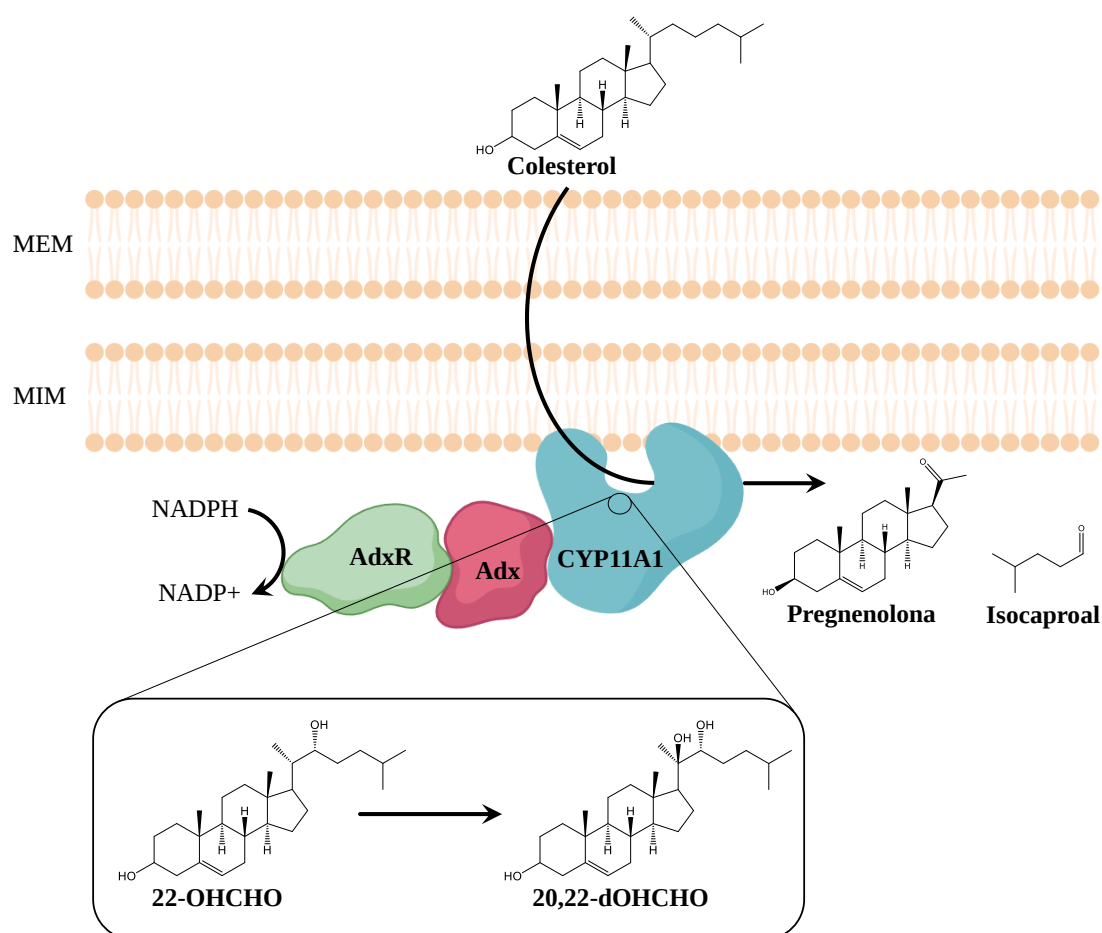


Figura I2. Esquema de la actividad colesterol hidroxilasa/liasa del CYP11A1. MEM: membrana externa mitocondrial; MIM: membrana interna mitocondrial; AdxR: adrenodoxina reductasa; Adx: adrenodoxina; 22-OHCHO: 22R-hidroxicolesterol; 20,22-dOHCHO: 20R,22R-dihidroxicolesterol.

1.3. Los esteroides en la industria farmacéutica

Los esteroides participan en una gran variedad de procesos fisiológicos y por tanto se pueden utilizar para el tratamiento y prevención de enfermedades relacionadas con la endocrinología, la reumatología, la inmunología o la oncología (Donova y Egorova, 2012). Distintos tipos de cáncer de colon, enfermedades neurodegenerativas, el shock anafiláctico, terapias de remplazamiento hormonal o la inhibición de la integrasa del VIH se tratan con fármacos esteroideos (Donova y Egorova, 2012). Este hecho ha incrementado el interés de la industria farmacéutica por los esteroides, de forma que actualmente constituyen el tipo de fármaco más vendido junto con los antibióticos, con unos ingresos de 3451 millones de USD en 2024 y con expectativas de seguir aumentando los próximos años (Donova, 2018; “Steroids Market size was USD 3451.2 million in 2024,” 2024).

Más de 300 fármacos esteroideos han sido aprobados para su uso en clínica y entre ellos podemos encontrar agentes antitumorales, antiinflamatorios, antimicrobianos, anticonceptivos, inmunosupresores, diuréticos o anabólicos (Donova, 2018). La Organización Mundial de la Salud en 2023 publicó una lista de fármacos que se consideran esenciales para la población y cuya disponibilidad y calidad debe estar asegurada. En esta lista podemos encontrar glucocorticoides como la dexametasona, usada para tratar el COVID-19, y progestágenos como la progesterona o la medroxiprogesterona, ampliamente utilizados como anticonceptivos o para el tratamiento de la amenorrea (“WHO Model List of Essential Medicines - 23rd list,” 2023). La mayoría de estos compuestos activos se producen mediante síntesis química, sin embargo, actualmente se conoce el potencial de las biotransformaciones microbianas en la síntesis de compuestos esteroideos, observándose una expansión en la aplicación de estos procesos (Donova y Egorova, 2012). En esta Tesis Doctoral se abordará así la biotransformación de esteroides para la producción microbiana de progesterona y de uno de sus precursores con interés industrial: el 22-hidroxi-23,24-bisnorcol-4-en-3-ona (4-HBC).

2. Biotecnología de los esteroides

La síntesis química de los esteroides necesita la aplicación de altas temperaturas y reactivos contaminantes para el medioambiente, además de tener bajos rendimientos y ocasionar ciertos problemas por la falta de estereoselectividad del proceso. Estos factores elevan el coste del proceso y hacen que requiera una mayor cantidad de sustrato, con un mayor consumo de recursos naturales (Donova y Egorova, 2012). Un ejemplo de este tipo de síntesis es la llamada degradación Marker. En esta reacción química, el sustrato de partida es la diosgenina, una sapogenina obtenida de las plantas que se somete a 3 pasos químicos con CrO_3 y 200 °C, para formar el acetato de 16-dehidropregnenolona (Hanson, 2005; Laveaga, 2005; Peng et al., 2021). Este intermediario, a continuación, se utiliza

para producir otros esteroides de interés. Por ejemplo, para la producción de progesterona, el acetato de 16-dehidropregnenolona se debe someter a 5 pasos más de síntesis química con un rendimiento del 50% (Fig. I3) (Peng et al., 2021).

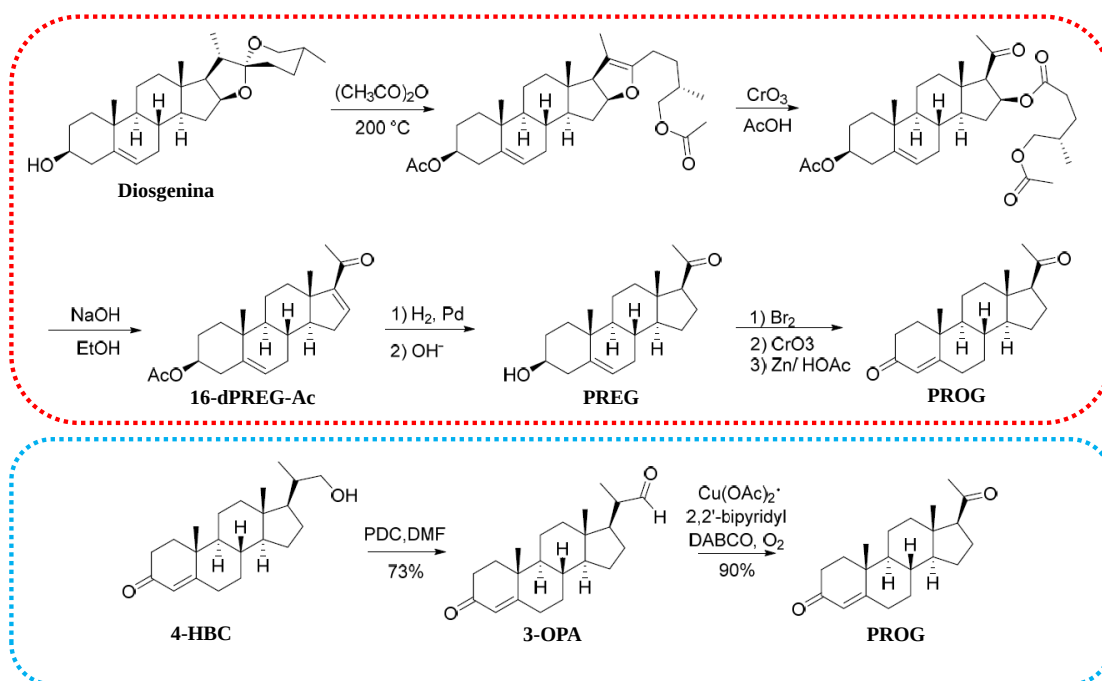


Figura I3. Síntesis química de progesterona a través de la degradación Marker (rojo) o a partir de intermediarios de biotransformación de esteroides (azul). 16-dPREG-Ac: acetato de 16-dehidropregnenolona; PREG: pregnenolona; PROG: progesterona; 4-HBC: 22-hidroxi-23,24-bisnorcol-4-en-3-ona; 3-OPA: (20S)-3-oxopregn-4-en-20-carboxaldehído. Adaptada de (Peng et al., 2021).

Las biotransformaciones bacterianas constituyen una importante herramienta para conseguir intermediarios esteroideos, ya que no solo permiten conseguir una funcionalización de la molécula de forma regio- y estereo-selectiva, sino que también permite completar varias reacciones en un único paso, reduciendo así la dependencia a los reactivos químicos y el empleo de condiciones menos extremas. Además, los microorganismos se pueden modificar para construir nuevas rutas metabólicas que permitan diseñar nuevos esteroides *à la carte* (Donova y Egorova, 2012). Un ejemplo que remarca la importancia de la aplicación de las biotransformaciones microbianas es la síntesis de cortisona. En este proceso se requerían 615 g de ácido desoxicólico, proveniente de la bilis de res, para producir tan solo 1 g de cortisona después de 31 pasos de síntesis química. Sin embargo, tras la introducción de unos pasos de biotransformación con *Rhizopus arrhizus* ATCC 11145 y *Aspergillus niger* ATCC 9142, el proceso se redujo a 11 pasos de síntesis química disminuyendo los costes de producción de 200 USD a 1 USD (Carballeira et al., 2009). Otra aproximación para la producción de esteroides que combina una semisíntesis química y microbiana es la producción de progesterona a partir de 4-HBC, que se obtiene mediante biotransformación microbiana de esteroides y después,

con tan solo 2 pasos de síntesis química, se obtiene la progesterona (Fig. I3) (Peng et al., 2021). Esta ruta de síntesis alternativa resulta más económica y respetuosa con el medioambiente.

2.1. Metabolismo microbiano de los esteroides

Para poder aplicar los microorganismos de una manera racional en los procesos biotecnológicos de biotransformación para la producción de esteroides, es de gran importancia conocer el metabolismo de los compuestos que se van a utilizar como sustratos y su regulación. El conocimiento actual de la degradación bacteriana de esteroides desarrollado sobre todo en los últimos años se detalla a continuación.

2.1.1. Transporte, catabolismo y regulación en actinobacterias

Las actinobacterias son un filo de bacterias Gram positivas caracterizadas por tener un alto contenido en GC en su genoma. Entre ellas podemos encontrar diferentes géneros como *Mycobacterium*, *Mycolicibacterium*, *Rhodococcus*, *Nocardia*, *Arthrobacter*, *Corynebacterium* o *Gordonia* y muchos de ellos presentan especies capaces de degradar total o parcialmente diversos esteroides (Arima et al., 1969; Brown y Peterson, 1966; Chipley et al., 1975; Ferreira y Tracey, 1984; Martin, 1977; Tak, 1942).

Tanto el transporte del colesterol al interior celular como su catabolismo y su regulación se han descrito con gran precisión sobre todo en las bacterias modelo *Mycobacterium tuberculosis* H37Rv, *Mycolicibacterium smegmatis* mc²155 y *Rhodococcus jostii* RHA1, principalmente porque este metabolismo es de gran importancia en la patogénesis de *M. tuberculosis* (Pandey y Sassetti, 2008). Estas bacterias se caracterizan por presentar una capa de ácidos micólicos en su envuelta celular y se ha propuesto que requieren de un transporte especializado que salve esta barrera. Esta actividad se consigue gracias a un sistema del tipo MCE (del inglés *mammalian cell entry*) que se encuentra codificado en el operón *mce4* descrito en *M. tuberculosis* H37Rv (Pandey y Sassetti, 2008), *M. smegmatis* mc² 155 (García-Fernández et al., 2017b) y *R. jostii* RHA1 (Mohn et al., 2008). En este operón se encuentran 10 genes que codifican transportadores del tipo ABC, junto con dos proteínas transmembrana que actúan como una permeasa y diversas proteínas encargadas de la unión y translocación del sustrato (García-Fernández et al., 2017b; Pandey y Sassetti, 2008). Estos genes se encuentran inducidos en presencia de colesterol y la mutación de algunos o de todos ellos impide el crecimiento en colesterol de *M. smegmatis* y *R. jostii* (García-Fernández et al., 2017b; Kendall et al., 2007; Mohn et al., 2008). En el caso de *M. tuberculosis*, en cambio, mutantes en el operón *mce4* solo disminuyen su velocidad de crecimiento en este esteroide, sugiriendo la actividad de un sistema de transporte alternativo desconocido en esta bacteria (Pandey y Sassetti, 2008).

Una vez el colesterol accede al interior celular, comienza su catabolismo aerobio descrito en la Figura I4. Este metabolismo se divide en dos eventos principales: la degradación de la cadena lateral y la apertura de los anillos del núcleo esteroideo. Comenzando con la degradación de la cadena lateral, el colesterol se transforma primero en colesteno al oxidarse el grupo hidroxilo del C3 e isomerizar el doble enlace. Esta reacción está catalizada por colesterol oxidasas (ChOx) o por 3 β -hidroxi- Δ^5 -esteroide deshidrogenasa/isomerasas (3 β -HSD), que es la actividad principal encontrada en *M. smegmatis* codificada por los genes *MSMEG_5228* y *MSMEG_5233* (Brzostek et al., 2013; Fernández de Las Heras et al., 2014; Uhía et al., 2011; Yang et al., 2007). Una vez obtenida la colesteno, o de manera simultánea, se requiere oxidar la cadena lateral para activarla y comenzar su degradación. Son dos CYP450, el CYP125A3 y el CYP142A2, los que oxidan el C27 dando lugar al ácido colest-4-en-3-ona-26-oico, que ya puede entrar en 3 ciclos similares a la β -oxidación de los ácidos grasos para metabolizar completamente la cadena lateral (García-Fernández et al., 2013; Ouellet et al., 2010; Rosłonec et al., 2009). En el primer ciclo, la cadena lateral se activa uniendo una molécula de CoA al nuevo grupo carboxilo por la enzima esteroide-CoA ligasa FadD19 (Wilbrink et al., 2011), para poder a continuación ser deshidrogenada por la deshidrogenasa ChsE4-ChsE5 y luego hidratada por la enoil-CoA hidratasa ChsH1-ChsH2, obteniendo finalmente 24-hidroxi-3-oxocolest-4-en-26-oil-CoA (Thomas et al., 2011; Thomas y Sampson, 2013; Yang et al., 2015, 2014). Esta nueva molécula es oxidada, aunque se desconoce la enzima involucrada en este paso. En *M. tuberculosis* se ha propuesto la deshidrogenasa codificada por el gen *Rv3502c*, generando 3,24-dioxocolest-4-en-26-oil-CoA, que es sustrato de la tiolasa FadA5 obteniendo así una molécula de propionil-CoA y 3-oxocolest-4-en-24-oil-CoA (Griffin et al., 2012; Nesbitt et al., 2010; Schaefer et al., 2015; Wiperman et al., 2013).

En el segundo ciclo vuelven a actuar una acil-CoA deshidrogenasa ChsE3 junto con una enoil-CoA hidratasa, todavía desconocida, obteniendo 22-hidroxi-3-oxo-colest-4-en-24-carboxil-CoA (22-OH-BCN-CoA). Este intermediario es sustrato de la β -hidroxiacil-CoA deshidrogenasa Hsd4A, codificada por el gen *MSMEG_5903* en *M. smegmatis*, y forma el compuesto 3,22-dioxo-colest-4-en-24-carboxil-CoA (22-OBNC-CoA) (Xu et al., 2016; Yang et al., 2015). Se ha propuesto que la tiolasa FadA5 vuelve a actuar y libera una molécula de acetil-CoA y genera 3-oxo-pregne-20-carboxil-CoA, acabando con el segundo ciclo de β -oxidación (Griffin et al., 2012; Nesbitt et al., 2010; Schaefer et al., 2015).

El tercer y último ciclo de β -oxidación es similar al primero, actuando la deshidrogenasa ChsE2-ChsE1 y la hidratasa ChsH2-ChsH1 y acaba con una rotura retroaldólica entre los últimos carbonos de la cadena lateral por la enzima Ltp2 (Thomas et al., 2011; Yang et al., 2015). Los productos finales obtenidos son otra molécula de propionil-CoA y el 17-cetoesteroide llamado 4-androsten-3,17-diona (AD), caracterizado por tener un grupo ceto en el C17 donde antes se encontraba la cadena lateral (Fig. I4). Este AD es un esteroide C19 de gran interés en la industria esteroidea y cuya importancia se desarrollará más adelante en el apartado 2.2.

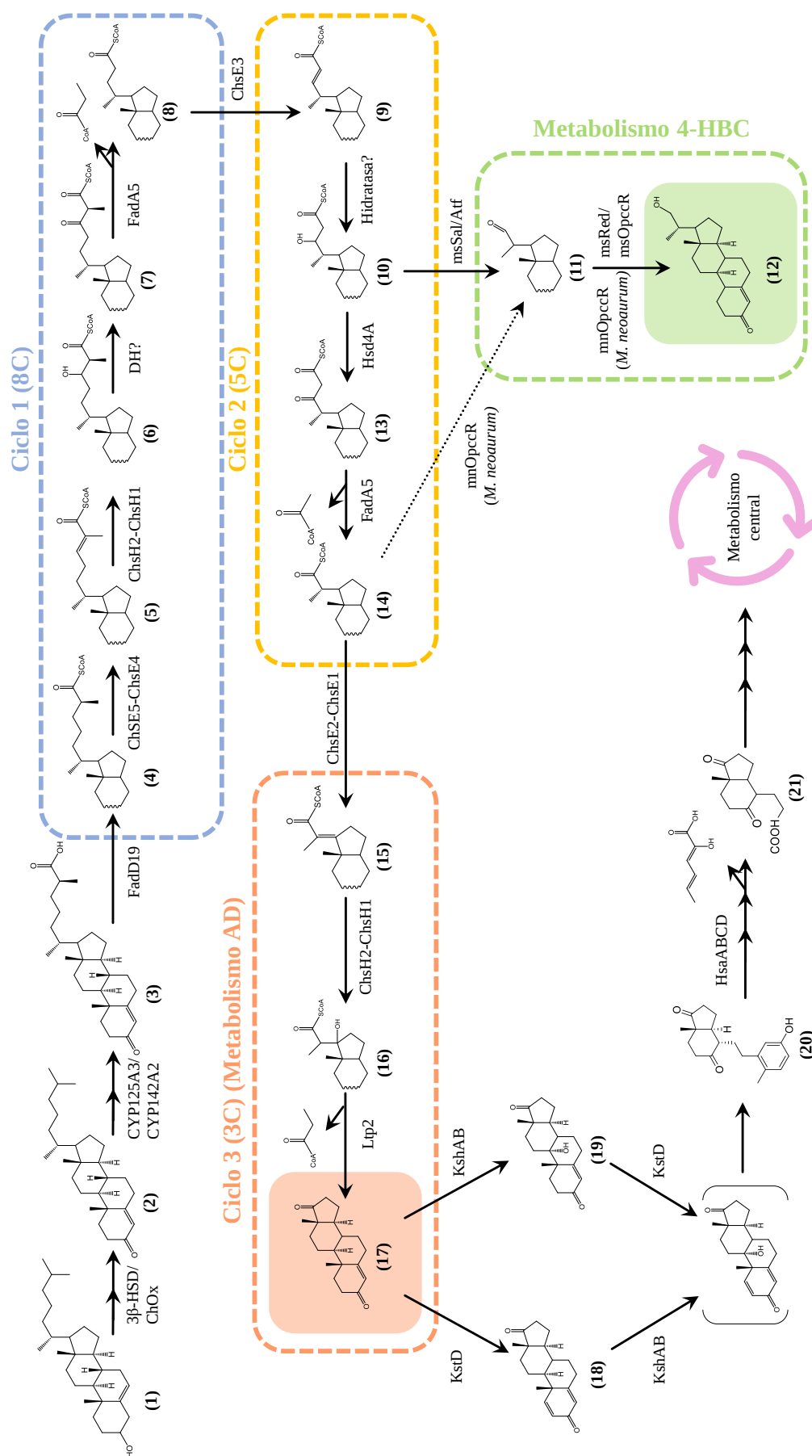


Figura I4. Metabolismo aeróbico del colesterol en actinobacterias. Los compuestos son: (1) colesterol; (2) colestenona; (3) ácido colest-4-en-3-ona-26-oico; (4) colest-4-en-3-ona-26-oil-CoA; (5) 3-oxocolest-4,24-dien-26-oil-CoA; (6) 24-hidroxi-3-oxocolest-4-en-26-oil-CoA; (7) 3,24-dioxocolest-4-en-26-oil-CoA; (8) 3-oxocolest-4-en-24-oil-CoA; (9) 3-oxocolest-4,22-dien-24-oil-CoA; (10) 22-hidroxi-3-oxo-colest-4-en-24-carboxil-CoA; (11) (20S)-3-oxopregn-4-en-20-carboxaldehído; (12) 22-hidroxi-23,24-bisnorcol-4-en-3-ona; (13) 3,22-dioxo-colest-4-en-24-carboxil-CoA; (14) 3-oxo-pregne-20-carboxil-CoA; (15) 3-oxo-4,17-pregnadiene-20-carboxil-CoA; (16) 17-hidroxi-3-oxo-4-pregnane-20-carboxil-CoA; (17) 4-androsten-3,17-diona; (18) 1,4-androstadien-3,17-diona; (19) 9 α -hidroxi-1,4-androstadien-3,17-diona; (20) 3-hidroxi-9,10-secoandrosta-1,3,5(10)-trien-9,17-diona; (21) dioxo-1,2,3,4,10,19-hexanorandrostan-5-oico. Las enzimas son 3 β -HSD: 3 β -hidroxi- Δ^5 -esteroide deshidrogenasa/isomerasa; ChOx: colesterol oxidasa; CYP125A3: citocromo P450 125A3; CYP142A2: citocromo P450 142A2; FadD19: esteroide-CoA ligasa; ChsE4-ChsE5: acil-CoA deshidrogenasa; ChsH2-ChsH1: enoil-CoA hidratasa; DH: deshidrogenasa; FadA5: tiolasa; ChsE3: acil-CoA deshidrogenasa; Hsd4A: β -hidroxiacil-CoA deshidrogenasa; msSal: aldolasa; msRed: reductasa; mnOpccR/msOpccR: acil-CoA reductasa; ChsE2-ChsE1: acil-CoA deshidrogenasa; Ltp2: acil-CoA liasa; KstD: 3-cetoesteroide- Δ^1 -deshidrogenasa; KshAB: cetoesteroide-9 α -hidroxilasa; HsaABCD: (di)oxigenasas.

Cabe destacar que en los últimos años se ha descrito una ruta alternativa en la degradación del colesterol en micobacterias, denominada ruta del 4-HBC, en la que la cadena lateral no se degrada por completo. Esta ruta empezaría en el 22-hidroxi-3-oxo-colest-4-en-24-carboxil-CoA (22-OH-BCN-CoA) y acabaría en el esteroide C22 4-HBC, que presenta una pequeña cadena alifática hidroxilada en C17 (Fig. I4). El 4-HBC es una sintona C22 de interés ya que se puede utilizar como punto de partida para la producción de progesterona. Al inicio de esta Tesis Doctoral, se desconocían las enzimas involucradas en este metabolismo. Peng et al. (2021) estudiaron la reducción del (20S)-3-oxopregn-4-en-20-carboxaldehído (3-OPA) a 4-HBC por la reductasa mnOpccR y propusieron que el 3-OPA se produce tras una eliminación retroaldólica del 22-OH-BCN-CoA. En esta Tesis Doctoral se describen la aldolasa msSal y la reductasa msRed involucradas en este metabolismo en *M. smegmatis*, encargadas de realizar la eliminación retroaldólica y la posterior reducción, respectivamente (Hernández-Fernández et al., 2024; Hernández-Fernández et al., 2023). Más adelante, Zhao et al. (2023) describieron en *Mycolicibacterium neoaurum* la aldolasa Atf homóloga a msSal.

La degradación aeróbica de los esteroides C19 continúa con una ruta metabólica común denominada ruta 9,10-seco y cuyo objetivo es la apertura de los anillos esteroideos (Fig. I4) (Bergstrand et al., 2016). El AD es deshidrogenado por una 3-cetoesteroide- Δ^1 -deshidrogenasa (KstD) para formar 1,4-androstadien-3,17-diona (ADD). En *M. smegmatis* hay 6 genes que codifican esta enzima (*kstD1*, 2, 3, 4, 5 y 6), sin embargo, solo *kstD1* (*MSMEG_5941*) actúa en el metabolismo del colesterol (Galán et al., 2017). Sobre el AD también actúa la cetoesteroide-9 α -hidroxilasa KshAB, una hidroxilasa de dos componentes que produce 9 α -hidroxi-1,4-androstadien-3,17-diona (9-OHAD). El componente oxigenasa de esta enzima es KshA, mientras que el componente reductasa es KshB. En *M. smegmatis* hay dos genes que codifican dos posibles KshA (*MSMEG_2870* y *MSMEG_5925*) y otros dos que codifican dos posibles KshB (*MSMEG_2893* y *MSMEG_6039*) (Galán et al., 2017). Cuando las dos actividades deshidrogenasa e

hidroxilasa actúan sobre la misma molécula, se genera un compuesto inestable que, de forma espontánea, abre el anillo B y se obtiene 3-hidroxi-9,10-secoandrosta-1,3,5(10)-trien-9,17-diona (3-HSA) (Fig. I4). Para abrir el anillo A de esta molécula, este se hidroxila primero por la enzima oxigenasa HsaAB y, posteriormente, por la meta-extradiol dioxigenasa HsaC se genera el ácido 4,5,9,10-diseco-3-hidroxi-5,9,17-trioxoandrosta-1(10),2-dien-4-oico (4,9-DSHA) (Dresen et al., 2010; Yam et al., 2009). Finalmente, este compuesto se hidroliza por la HsaD y se obtiene 2-hidroxihexa-2,4-dienoico (HHD) y dioxo-1,2,3,4,10,19-hexanorandrostano-5-oico (HIP) (Lack et al., 2008, 2010). Ambos compuestos acaban entrando en las rutas centrales del metabolismo bacteriano degradándose por completo el esqueleto central esteroideo (Horinouchi et al., 2005; Kieslich, 1985; van der Geize et al., 2011).

La expresión de todas las proteínas involucradas en el transporte y las enzimas catabólicas del colesterol requiere una gran cantidad de energía, de forma que el catabolismo del colesterol se encuentra altamente regulado por los represores transcripcionales KstR y KstR2. Estos reguladores son unos represores de la familia TetR y se encuentran muy bien estudiados en *M. smegmatis*, *M. tuberculosis* y *R. jostii* RHA1 (Kendall et al., 2007; Uhía et al., 2011). El represor KstR se encarga de controlar la expresión del regulón *kstR* en donde se encuentran 83 genes involucrados en el transporte y la ruta alta y central de la degradación del colesterol. Su secuencia operadora se ha identificado como TnnAACnnGTTnnA y se ha comprobado que su inductor es el ácido 4-colesten-3-ona-26-oil-CoA (García-Fernández et al., 2014; Ho et al., 2016; Uhía et al., 2011). El regulón *kstR2* a su vez está regulado por el represor KstR2 y controla la expresión de 15 genes catabólicos implicados en la ruta baja del colesterol. Se sabe que el HIP-CoA es el inductor del represor (Casabon et al., 2013) y su secuencia operadora es AAnCAAGnnCTTGnTT (García-Fernández et al., 2015). Recientemente se ha descrito un tercer regulador en *M. neoaurum*, denominado KstR3, que aparentemente reprime la expresión del *cluster* HBC involucrado en el metabolismo del 4-HBC, sin embargo, se desconoce tanto su inductor como la región operadora (Zhao et al., 2023).

2.1.2. Metabolismo del colesterol en bacterias Gram negativas

La ruta catabólica del colesterol descrita en actinobacterias no se encuentra en su totalidad presente en las bacterias Gram negativas capaces de degradar esteroides. Las primeras reacciones observadas en este tipo de bacterias son la transformación anaerobia de colesterol a coprostanol por bacterias intestinales (Freier et al., 1994; Li et al., 1995), aunque ninguna de ellas era capaz de mineralizar el colesterol o el coprostanol. Otra modificación del colesterol que se ha descrito es la que llevan a cabo dos CYP450scc de *Novosphingobium aromaticivorans* y *Novosphingobium subterraneum* (Yamagishi et al., 2011). En este caso, ambos CYP450scc son capaces de cortar la cadena lateral del colesterol y obtener así pregnenolona, de igual forma que sucede con el CYP11A1 de mamíferos. Esta actividad CH/L sería la primera detectada en CYP450scc bacterianos. Sin embargo, esta reacción solo se ha observado *in vitro* y la capacidad de estas especies de degradar colesterol no se ha comprobado.

Por el momento, solo se han descrito dos bacterias Gram negativas, pertenecientes al grupo de las β -proteobacterias, capaces de crecer en presencia de colesterol como única fuente de carbono y de energía, tanto en condiciones anaerobias como aerobias: *Sterolibacterium denitrificans* DSM13999 y 72Chol DSM12783 (no clasificada) (Harder y Probian, 1997; Lin et al., 2015; Tarlera y Denner, 2003; Wang et al., 2013a).

En este metabolismo, esquematizado en la Figura I5, el colesterol se oxida y se isomeriza a colesteno por la colesterol deshidrogenasa/isomerasa AcmA que, a continuación, se oxida a colest-1,4-dien-3-ona por la colesteno- Δ^1 -deshidrogenasa AcmB (Chiang et al., 2008b, 2008a). Tras 4 pasos de hidroxilaciones y deshidrogenaciones para activar la cadena lateral, se obtiene desmost-1,4-diene-3-ona-CoA (Jacoby et al., 2020; Warnke et al., 2017). A partir de aquí comienza la degradación de la cadena lateral de una forma similar a la observada en actinobacterias en la que en cada ciclo de β -oxidación actúan diferentes CoA ligasas, deshidrogenasas, hidratasas y aldolasas hasta acabar con un grupo ceto en C17 formando ADD (Wang et al., 2013b; Warnke et al., 2018, 2017). Este intermediario entra ahora en la ruta 2,3-*seco*, en vez de la previamente descrita 9,10-*seco*. Esta ruta es utilizada también en la degradación de andrógenos como la testosterona y se caracteriza por ser un metabolismo anaerobio, en el que primero el anillo A se abre tras una serie de hidroxilaciones donde el agua actúa como donador de átomos de oxígeno y se forma 1,17-dioxo-2,3-secoandrostan-3-oico (DSAO) (Chiang et al., 2008b; Wang et al., 2013b). Se desconocen las siguientes reacciones que continúan con la mineralización de DSAO, pero sí que se ha detectado HIP como intermediario de esta ruta, por lo que se propone que entraría al metabolismo central de igual forma que se ha observado en la ruta 9,10-*seco* (Warnke et al., 2018, 2017).

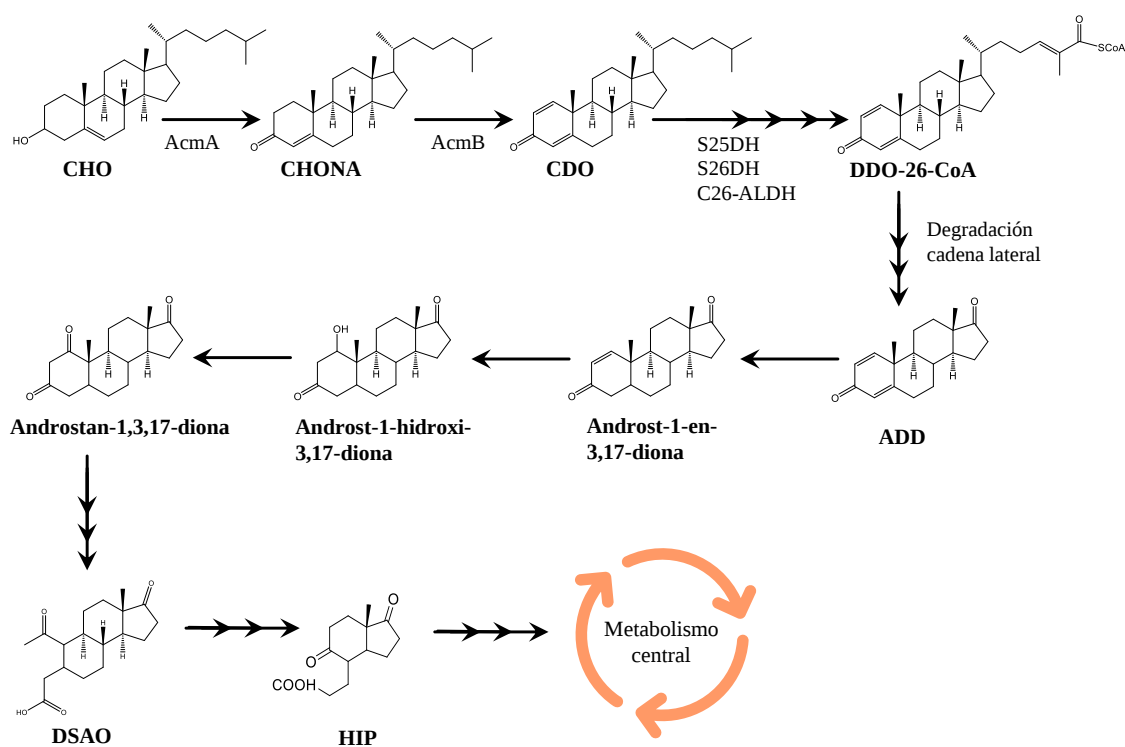


Figura I5. Ruta de degradación anaerobia del colesterol de *S. denitrificans*. CHO: colesterol; CHONA: colesteno; CDO: colest-1,4-dien-3-ona; DDO-26-CoA: derivado CoA del ácido DDO-26-oico; ADD: 1,4-androstadien-3,17-diona; DSAO: ácido 1,17-dioxo-2,3-seco-androstan-3-oico; HIP: dioxo-1,2,3,4,10,19-hexanorandrostan-5-oico; AcmA: colesterol deshidrogenasa/isomerasa; AcMB: colesteno- Δ^1 -deshidrogenasa; S25DH: esteroide C-25 deshidrogenasa; S26DH: esteroide C-26 deshidrogenasa; C26-ALDH: C-26 aldehído deshidrogenasa.

2.2. Producción microbiana de sintonas esteroideas

Estos estudios metabólicos permiten conocer cuáles son los intermediarios esteroideos que son de potencial interés industrial. Este tipo de moléculas se denominan sintonas. El uso de cepas silvestres no suele dar buen resultado a la hora de producir estas sintonas ya que se obtiene una mezcla de productos o no se evita la completa degradación del núcleo de los esteroides (Donova y Egorova, 2012). Por eso, utilizar las herramientas genéticas desarrolladas para cada tipo de bacteria permite obtener distintos mutantes que acumulan las sintonas deseadas.

Para la producción industrial de esteroides, se han empleado las sapogeninas como materia prima durante décadas, como es el caso de la degradación Marker descrita anteriormente. El uso de este sustrato genera algunos problemas, como el agotamiento de recursos vegetales silvestres o altos costes (Tong y Dong, 2009; Wang et al., 2011). Por otro lado, los controles de calidad requeridos para utilizar cualquier materia prima de origen animal dificultan el uso del colesterol, obtenido de aceites y grasas animales, como sustituto de las sapogeninas (Fernandes et al., 2003; Malaviya y Gomes, 2008). Es por eso que los fitoesteroides se plantean como una verdadera alternativa más barata, empleándose actualmente como sustrato. Los fitoesteroides pueden ser biotransformados por distintas bacterias y se pueden obtener de desechos agrícolas o industriales, como por ejemplo residuos de la industria papelera o de la soja, lo que incluye este proceso en una economía circular más sostenible (Donova, 2021).

La producción de las sintonas esteroideas C19, como el AD, o las C22, como el 4-HBC, son el principal objetivo de la industria de esteroides, ya que sirven como punto de partida para sintetizar una gran diversidad de fármacos (Fig. I6). Estas sintonas se pueden conseguir biotransformando fitoesteroides para luego, mediante procesos de síntesis química, ser modificados. Así, el AD se utiliza como precursor de andrógenos, estrógenos o glucocorticoides y el 4-HBC se utiliza para la síntesis de progestágenos (Donova, 2021). La producción de AD, ADD o 9-OHAD mediante biotransformaciones ya se ha implementado a escala industrial en varios países como España, EEUU, Alemania, China e India, y el mercado de AD/ADD está valorado en 1000 millones de USD anuales (Donova, 2021; Donova y Egorova, 2012).

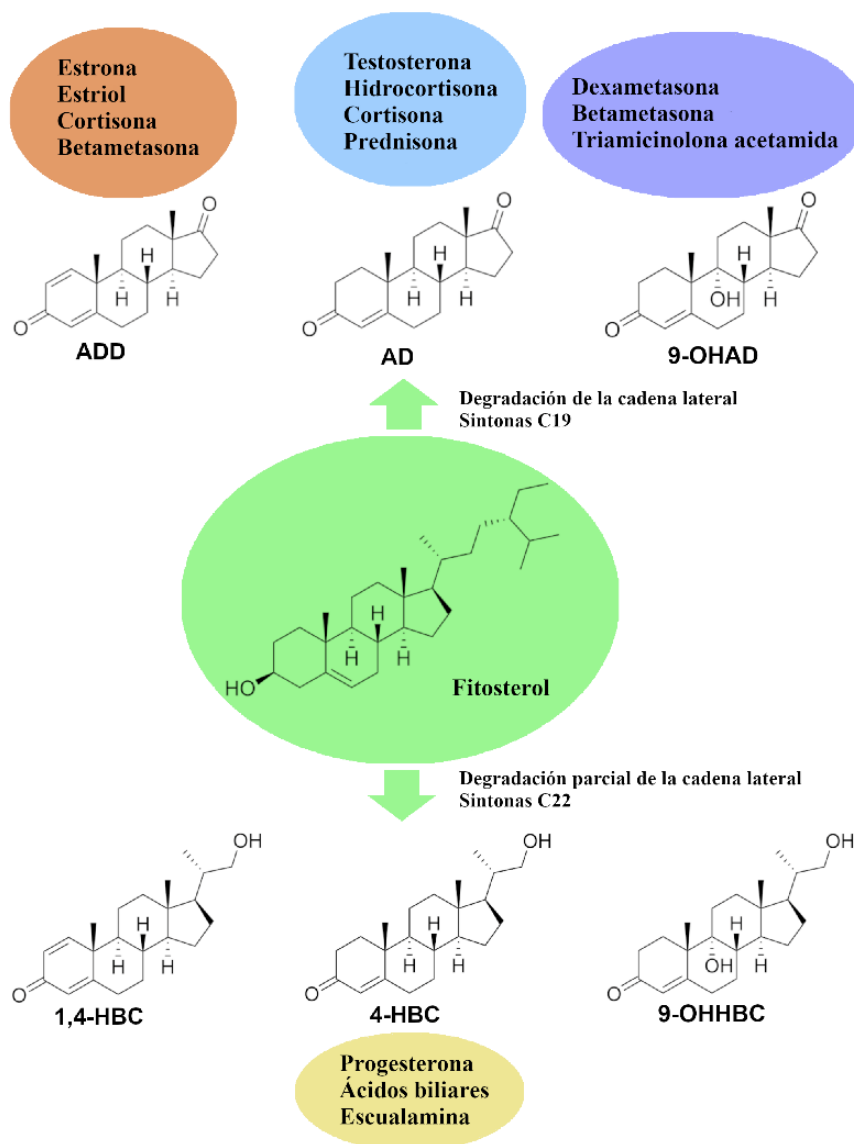


Figura I6. Estructura de las sintonas C19 y C22 y los fármacos esteroides que se pueden sintetizar con ellas. AD: 4-androsten-3,17-diona; ADD: 1,4-androstadien-3,17-diona; 9-OHAD: 9 α -hidroxi-1,4-androstadien-3,17-diona; 4-HBC: 22-hidroxi-23,24-bisnorcol-4-en-3-ona; 1,4-HBC: 22-hidroxi-23,24-bisnorcol-1,4-dien-3-ona; 9-OHHBC: 9,22-dihidroxi-23,24-bisnorcol-4-en-3-ona.

Sin embargo, uno de los problemas a los que se enfrenta la biotransformación de esteroides es la actividad de un metabolismo secundario en las cepas empleadas como factorías celulares. Como se ha descrito antes, en la degradación del colesterol el intermediario principal es el AD, pero parte del catabolismo se desvía a la producción de 4-HBC, de forma que se ha observado que en las bacterias productoras de AD se acumula 4-HBC como subproducto. Del mismo modo, al diseñar factorías celulares productoras de 4-HBC, se ha observado que éstas acumulan cierta cantidad de AD. Esta situación no solo reduce el rendimiento del proceso, sino que además dificulta la purificación del compuesto de interés, elevando los costes de producción e impidiendo su aplicación industrial (Galán et al., 2017; Han et al., 2024; Peng et al., 2021).

Otro ejemplo que reitera la necesidad de diseñar cepas industriales optimizadas es la producción de hidrocortisona. Los hongos filamentosos no son capaces de utilizar los esteroides como fuente de carbono y de energía, pero sí pueden modificarlos de forma que, para la producción de hidrocortisona, se emplea el hongo *Curvularia lunata*, que contiene un CYP450 que realiza el paso final de 11 β -hidroxilación. Sin embargo, el empleo de la cepa silvestre implica otras actividades hidroxilasas que generan esteroides di-, tri- y polihidroxilados, lo que complica los procesos de purificación (Donova, 2021; García-Fernández et al., 2017a).

La aplicación de la biotecnología ha permitido desarrollar diferentes métodos biocatalíticos, así como conseguir nuevas cepas que permiten realizar bioprocesos mejorados que no solo proporcionen tanto compuestos conocidos como nuevos, sino que se obtengan de una forma mucho más eficaz y selectiva. En micobacterias, la mutación de las enzimas KstD, KshAB o ambas permite la acumulación de 9-OHAD, ADD o AD, respectivamente. Así, la cepa MS6039-5941 de *M. smegmatis*, mutante en ambas enzimas, acumula AD a partir de fitoesteroides, y la expresión heteróloga de una 17 β -hidroxiesteroide deshidrogenasa (17 β -HSD) permite la producción de testosterona (Fernández-Cabezón et al., 2018a; Galán et al., 2017). La producción de boldenona, una versión deshidrogenada de la testosterona, se ha conseguido también mediante dos pasos obteniendo primero AD gracias al uso de micobacterias y su posterior Δ^1 -deshidrogenación por el hongo *Fusarium* sp. (Kutney et al., 2003). Una cepa de *M. neoaurum* ATCC 25795 triple mutante de tres enzimas KstDs que además sobreexpresa KshB produce de manera eficaz 9-OHAD (Wei et al., 2014). Estas cepas productoras de sintonas C19 presentan el problema expuesto previamente y es que todas acumulan como subproducto esteroides C22. Las sintonas C19 no sirven para sintetizar progesterona, pero como su metabolismo está ligado al de los esteroides C22 y debido a que al inicio de esta Tesis Doctoral no se había descrito en la literatura una micobacteria recombinante capaz de producir AD de forma pura sin la presencia de otros intermediarios, se estableció como uno de los objetivos a conseguir en este trabajo.

Para construir cepas productoras de esteroides C22, se genera la delección de la deshidrogenasa Hsd4A que permite desviar el flujo del metabolismo hacia la producción de 4-HBC ya que así la cadena lateral de su precursor, el 22-OH-BCN-CoA, no se degrada por completo (Fig. I4). En concreto, la mutación de Hsd4A en una cepa de *Mycolicibacterium fortuitum* productora de AD permitió la acumulación altas concentraciones de 4-HBC (Liu et al., 2023a). Del mismo modo, cuando el mutante Δ kstD-hsd4A de *M. neoaurum* sobreexpresó además la reductasa MnOpccR, la ratio molar de 4-HBC aumentó hasta el 93% (Peng et al., 2021). Al inicio de esta Tesis Doctoral, no existía ningún mutante de *M. smegmatis* que produjera 4-HBC y tanto su metabolismo como los genes implicados en él permanecían desconocidos.

La progesterona es considerada un esteroide clave en la industria farmacéutica. Su producción a partir de 4-HBC ahorra varios pasos de síntesis química y permite reducir costes al usar fitoesteroides como materia prima, pero en este proceso todavía se requieren reactivos químicos que van en contra de conseguir una industria más respetuosa con el

medio ambiente. Con el objetivo de conseguir una completa producción biológica de progesterona, se han llevado a cabo distintas aproximaciones en las que el sistema CH/L de mamíferos se ha expresado de forma heteróloga en distintos microorganismos, como *Saccharomyces cerevisiae* (Duport et al., 1998), *E. coli* (Makeeva et al., 2013), *Bacillus megaterium* (Gerber et al., 2015), *Yarrowia lipolytica* (Zhang et al., 2019), y *M. neoaurum* (Liu et al., 2022). La cantidad de progesterona producida por estas cepas recombinantes se encuentra entre 0,4 y 235 mg/L, en función de la factoría celular utilizada, el sistema CH/L expresado y el sustrato de partida utilizado (colesterol, fitoesteres o intermediarios del ergosterol). Estos niveles son todavía muy inferiores a los requeridos para su aplicación industrial, además de ser necesario un perfeccionamiento de las factorías celulares, ya que los rendimientos de purificación son más bajos por la acumulación de intermediarios como 4-HBC o ésteres de pregnenolona (Duport et al., 1998; Liu et al., 2022; Zhang et al., 2019). Al inicio de esta Tesis Doctoral, no se había estudiado el uso de *M. smegmatis* como biofactoría para la producción de este esteroide, por lo que se planteó como uno de los objetivos principales de este trabajo. Sin embargo, recientemente, Karpov et al. (2024) describieron la producción de progesterona en una cepa de *M. smegmatis* que expresa el CYP11A, la Adx y la AdxR bovinas, llegando a acumular casi 20 mg/L de progesterona y 450 mg/L de MOM-pregnenolona. Estos datos, junto con los resultados obtenidos en el Capítulo 1, serán discutidos más adelante.

Cabe destacar que la progesterona no solo tiene aplicación clínica por sí misma, sino que también se puede usar como punto de partida para la síntesis de otros compuestos esteroideos como la hidrocortisona (Pan et al., 2023). En este caso, se transformó la progesterona en hidrocortisona en *Escherichia coli* expresando tres CYP450 de mamífero en cascada (la 17 α -hidroxilasa CYP17A1, la 21-hidroxilasa CYP21A2 y la 11 β -hidroxilasa CYP11B1), llegando a producir casi 17 μ M de hidrocortisona en 24 h.

3. Los esteroides como contaminante medioambiental

La producción industrial de esteroides tanto naturales como sintéticos y su liberación en el medioambiente puede generar problemas ambientales significativos. Es por eso que un campo importante de estudio se centra en conocer el destino, el impacto y cómo se eliminan del medioambiente los esteroides que implican un riesgo para la salud de los nichos ecológicos. Los disruptores endocrinos (EDC, del inglés *endocrine-disrupting chemicals*) son un grupo de compuestos químicos que incluyen fitosanitarios, pesticidas, productos de uso industrial y, especialmente, hormonas tanto naturales como sintéticas (Zhao et al., 2018). Estos EDCs se caracterizan por actuar a nivel sistémico interfiriendo con el sistema endocrino de animales y humanos en diferentes etapas de su desarrollo y a bajas concentraciones, de forma que supone un riesgo para la salud (Vandenberg et al., 2012).

El incremento de la población mundial y animal de granja ha aumentado la producción de esteroides y su consumo como fármacos, lo que desemboca a su vez en un

aumento de la presencia de estos EDCs en los ecosistemas (Lange et al., 2002). Una mujer embarazada puede llegar a excretar hasta 7 mg al día de estrógenos, mientras que la excreción anual de este tipo de hormonas por parte de los animales de granja de la Unión Europea y EEUU alcanza hasta 49 T (Johnson et al., 2000; Lange et al., 2002). En las aguas residuales de una fábrica farmacéutica de Pekín se han llegado a detectar hasta 4,4 µg/L de estrógenos como estrona (E1), estradiol (E2), estriol (E3), etinilestradiol (EE2), dietilestilbestrol (DES) y estradiol-17-valerato (Zhou et al., 2012) mientras que, en un estudio sobre la actividad estrogénica en efluentes de 75 plantas depuradoras de agua europeas, en 27 de ellas se detectaron niveles superiores a 0,5 ng/L de equivalentes de estradiol (EEQ), alcanzando un máximo de 17,9 ng/L EEQ (Jarošová et al., 2014). La progesterona juega un papel esencial en humanos por su implicación en el ciclo menstrual y el embarazo y es ampliamente utilizada como anticonceptivo oral junto con el EE2, además de su uso en terapias de reemplazamiento hormonal o en la medicina veterinaria (Kolatorova et al., 2022). Por ello, esta hormona también se ha detectado en aguas residuales a concentraciones entre 0,11 y 439 ng/L e incluso en aguas superficiales hasta 1,1 ng/L (Golovko et al., 2018; Kasambala et al., 2019). De hecho, aunque emplear estanques de estabilización reduce en un 75% los niveles de este EDC, la concentración detectada en el efluente todavía supone un riesgo para los organismos acuáticos que viven en los ecosistemas donde estas aguas son vertidas (Kasambala et al., 2019).

Las hormonas sexuales por tanto llegan al medio ambiente a través de estos efluentes de plantas depuradoras de agua o por la aplicación agrícola del estiércol de ganado como fertilizante (Hanselman et al., 2003; Lorenzen et al., 2004) y terminan contaminando abonos, aguas superficiales, acuíferos, tierras de cultivo, sedimentos de estuarios y el agua potable (Fernández et al., 2017; Griffith et al., 2016; Yin et al., 2002). Una vez en los ecosistemas, estos esteroides se consideran compuestos muy recalcitrantes, debido a la falta de grupos funcionales en su estructura y su baja solubilidad en agua, lo que supone un riesgo para la fauna presente en este nicho ecológico. Cuando peces de la especie *Pimephales promelas* son expuestos a progesterona, se observa una reducción en la expresión de vitelogenina (una proteína utilizada como indicador de feminización en animales) y del receptor de andrógenos, lo que señala la actividad desfeminizante de los progestágenos (Sangster et al., 2016). En esta especie de peces se ha estudiado también el impacto de hormonas feminizantes como los estrógenos y se ha observado que estos EDCs afectan a la supervivencia y la fecundidad de los individuos (Schwindt et al., 2014). Se ha demostrado incluso que una exposición crónica a EE2 provoca el colapso de la población de estos peces (Kidd et al., 2007). El impacto de estas hormonas también alcanza a animales invertebrados como moluscos, corales o rotíferos (Scott, 2012; Stout et al., 2010; Tarrant et al., 2004). Por ejemplo, la exposición a E2 por parte de larvas de *Austrominius modestus* no solo aumenta la acumulación de mutaciones en su ADN, sino que también aumenta sus niveles de vitelogenina (Atienzar et al., 2002; Billingham et al., 2000). El E2 es capaz también de provocar una disminución en la cantidad de paquetes de esperma y de huevos liberados por parte de los corales, además de disminuir su tasa de crecimiento esquelético (Tarrant et al., 2004).

Hay varios estudios que apuntan al potencial carcinógeno de los EDCs (Alsen et al., 2021; Fucic et al., 2012; Pal et al., 2025) e incluso los estrógenos y progestágenos se encuentran incluidos en el Grupo 1 de compuestos peligrosos para la salud humana por la Agencia Internacional para la Investigación del Cáncer de la OMS (“Agents Classified by the IARC Monographs, Volumes 1–138 – IARC Monographs on the Identification of Carcinogenic Hazards to Humans,” 2025). Los mecanismos a través de los cuales los EDCs estrogénicos contribuyen al desarrollo del cáncer de mama, por ejemplo, son todavía desconocidos. Se piensa que estos compuestos desregulan la señalización celular y alteran el metabolismo, ejerciendo así como promotores de la proliferación de células neoplásicas (Winz et al., 2023). El plastificante bisfenol A (BPA) se ha demostrado que aumenta los niveles de expresión del canal de potasio voltaje-dependiente Kv3.4, lo que induce la migración celular en el modelo de cáncer de mama MCF7 (Sim et al., 2025).

3.1. Degradación de esteroides con actividad EDC

La correcta eliminación de los EDCs en las aguas residuales o desechos industriales y agrarios es clave para evitar la contaminación de los ecosistemas y los efectos de salud que desencadenan. Los tratamientos utilizados de forma tradicional en la depuración de aguas como la coagulación con cationes de aluminio y hierro o la cloración no son lo suficientemente efectivos para acabar con los EDCs. Otros tratamientos como la oxidación química, la ozonización, el carbono activo o el uso de membranas son alternativas más eficaces, siendo la adsorción el método más competente a la hora de eliminar estos compuestos de residuos industriales. Esto se debe a que es el sistema más simple, con un manejo muy fácil y además evita la formación de otros compuestos secundarios peligrosos (Azizi et al., 2022; Waleed et al., 2024).

Otra opción eficaz para eliminar EDCs de aguas residuales y desechos consiste en la biodegradación microbiana. Estos microorganismos pueden mineralizar completamente los compuestos o modificarlos de forma que pierdan su actividad EDC. Existen numerosas bacterias pertenecientes a los filos o clases Actinobacteria, Firmicutes o Proteobacteria capaces de degradar distintos tipos de EDC y, especialmente, hormonas sexuales (Arai et al., 1998; Balázs et al., 2014; Ding et al., 2016; Fernández et al., 2017; Kornman y Loesche, 1982; Qin et al., 2016; Yoshimoto et al., 2004). Los microorganismos son un componente fundamental de los fangos activos de las depuradoras y, gracias a este método, se puede conseguir entre el 90-97% de degradación del BPA y del triclosán (Zhou et al., 2015). Respecto a las hormonas sexuales, estas llegan a las depuradoras en forma de conjugados glucurónidos o sulfónidos y luego son desconjugados por actividades β -glucuronidasas y arilsulfatasas, liberándose así hormonas como E1, E2, E3 y EE2 (Azizi et al., 2022). La eliminación de estas hormonas tanto por la actividad microbiana de los fangos activos como por su adsorción a los lodos oscila entre el 61 y el 85% (Azizi et al., 2022). Existen otras técnicas que se fundamentan en la biodegradación de los EDCs por microorganismos como son las lagunas aireadas o los biorreactores de membrana. Estas lagunas son un método más convencional que se aplican en áreas rurales para tratar aguas residuales domésticas y se ha comprobado que

no solo disminuyen los niveles de estrógenos hasta no ser detectables, sino que eliminan entre un 88% y un 100% otros contaminantes (Li et al., 2013). Los biorreactores de membrana constituyen un método prometedor para el tratamiento de aguas residuales ya que proporcionan una mayor flexibilidad a la hora de trabajar con mayores tiempos de retención. Estos biorreactores permiten eliminar más del 90% de los esteroides junto con procesos de nitrificación y desnitrificación y pueden conseguir una mayor eficiencia en la eliminación de EDCs debido a que acumulan un mayor número de microorganismos degradadores al aumentar su retención en los fangos (Li et al., 2013).

3.1.1. Degradación aeróbica de andrógenos

El estudio en detalle del metabolismo de los EDCs en microorganismos nos permite poder desarrollar aplicaciones, ya sea para descontaminar o para construir biosensores y detectar la presencia de estos compuestos en los ecosistemas. La testosterona es una hormona sexual androgénica cuya mineralización aeróbica se ha descrito de forma prácticamente completa en la β -proteobacteria *Comamonas testosteroni* TA441 mediante modificación genética (Horinouchi et al., 2019, 2012, 2004b, 2004a, 2003, 2001). Recientemente, este metabolismo también se ha estudiado en la α -proteobacteria *Caenibius tardaugens* NBRC 16725, aislada de los fangos activos de una depuradora de agua de Tokio y que utiliza estrógenos como única fuente de carbono y de energía (Fujii et al., 2003).

En este metabolismo, el 17 β -hidroxilo de la testosterona se oxida a un grupo ceto por la 17 β -HSD y se forma AD. A continuación, de igual forma que ocurre en el catabolismo del colesterol, este esteroide C19 se degrada por la ruta 9,10-seco, primero deshidrogenándose a ADD por las deshidrogenasas KstD y luego hidroxilándose por las hidroxilasas KshAB. Cuando ambas actividades ocurren a la vez, se abre el anillo B espontáneamente y se aromatiza el anillo A. Este nuevo seco-esteroide es el 3-HSA que, tras una serie de hidroxilaciones, abre el anillo A y se forma el intermediario HIP, cuyo destino final es entrar en el metabolismo central de la célula a través de la ruta baja.

Todas estas actividades enzimáticas están codificadas en dos *clusters* génicos de 10 y 18 genes catabólicos descritos en *C. testosteroni* TA441 y están presentes en otras proteobacterias degradadoras de andrógenos como *Burkholderia*, *Caenibius*, *Cupriavidus*, *Novosphingobium*, *Pseudoalteromonas*, *Pseudomonas*, *Shewanella* y *Sphingomonas* (Bergstrand et al., 2016; Horinouchi et al., 2012). Este *cluster* se encuentra regulado por TeiR, un activador transcripcional de la familia de LuxR cuya mutación impide el crecimiento de *C. testosteroni* en testosterona (Horinouchi et al., 2004b; Pruneda-Paz et al., 2004). En *C. tardaugens* NBRC 16725 solo se encuentra un *cluster* SD (del inglés *steroid degrading*) responsable de la degradación de la testosterona y otros esteroides C19 de esta bacteria. Esta agrupación de genes se expresa de forma basal y se compone de dos regiones de 5 y 21 genes que se transcriben de forma divergente (Ibero et al., 2019).

Las bacterias capaces de degradar el colesterol, como las del género *Mycobacterium*, *Mycolicibacterium*, *Rhodococcus* y *Gordonia*, presentan otro set de enzimas específico para degradar esteroides C19, como el andrógeno AD, y que difiere del empleado para metabolizar el colesterol (Fernández de las Heras et al., 2012; Fernández-Cabezón et al., 2018b; Li et al., 2014; Petrusma et al., 2011; van Der Geize et al., 2000). En *M. smegmatis*, por ejemplo, Fernández-Cabezón et al. (2018b) describieron el *cluster* C-19+, en el que se encuentran varias isoenzimas de KstD (KstD2 y KstD3), KshAB (KshA2 y KshB2), HsaA (HsaA2), HsaC (HsaC2) y HsaD (HsaD2) y cuya expresión se encuentra reprimida por el regulador transcripcional PadR. Estas isoenzimas catalizarían también la deshidrogenación y la hidroxilación del AD para abrir la estructura central de esterano y formar 3-HSA y, posteriormente, HIP. Este *cluster* C-19+ habría evolucionado independientemente del regulón *kstR*, involucrado en la transformación del colesterol a HIP, que después se mineralizaría por los genes codificados en el regulón *kstR2*.

3.1.2. Degradación aeróbica de progestágenos

A pesar de que se han descrito numerosos microorganismos capaces de metabolizar la progesterona como bacterias de los géneros *Ballicus*, *Comamonas*, *Rhodococcus*, *Enterobacter* y *Pseudomonas* (Fernández de Las Heras et al., 2009; Horinouchi et al., 2010; Liu et al., 2013, 2020), no hay información relativa a los genes implicados en los primeros pasos de la mineralización de esta hormona. Mediante análisis de intermediarios utilizando bacterias aisladas de aguas residuales, se han detectado en el medio de cultivo compuestos como AD, ADD, testosterona, boldenona, 1,2-dehidroprogesterona, 6,7-dehidroprogesterona, 17 α -hidroxiprogesterona y alopregnalona (Fig. I7) (Liu et al., 2013, 2020). Horinouchi et al. (2012) sugieren que la degradación de progesterona en *C. testosteroni* se lleva a cabo por la ruta 9,10-*seco*, pero las enzimas que canalizarían el catabolismo de la hormona hacia esta ruta no están descritas.

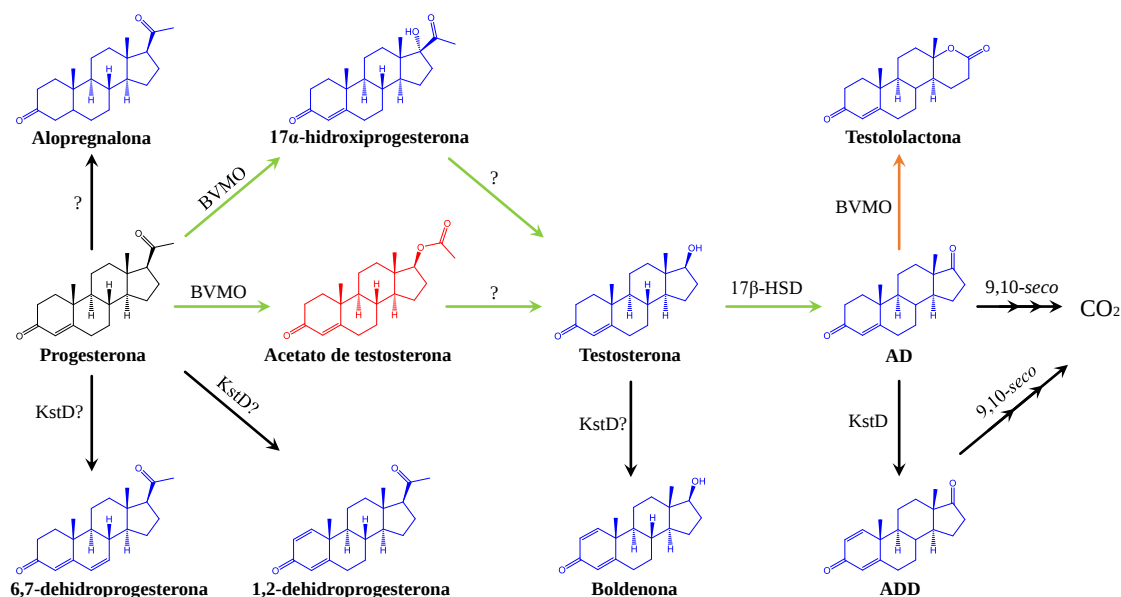


Figura I7. Rutas metabólicas propuestas para la degradación microbiana de la progesterona. Las moléculas en azul son compuestos detectados *in vivo* en la bibliografía, mientras que las marcadas en rojo se han detectado *in vitro*. Las flechas negras corresponden a actividad enzimática observada en bacterias, las flechas naranjas a actividad observada en hongos y las flechas verdes a actividades enzimáticas observadas en bacterias y hongos. AD: 4-androsten-3,17-diona; ADD: 1,4-androstadien-3,17-diona; BVMO: Baeyer-Villiger monooxigenasa; KstD: 3-cetoesteroide- Δ^1 -deshidrogenasa; 17 β -HSD: 17 β -hidroxiesteroide deshidrogenasa.

Algunas especies de hongos pueden modificar la cadena lateral de la progesterona transformándola en otros andrógenos como la testololactona, que presenta actividad anticancerígena, mediante varios pasos catalizados por Baeyer-Villiger monooxigenasas (BVMOs). Javid et al. (2018) proponen que la BVMO de *Aspergillus sojae* PTCC 5196 introduce un átomo de oxígeno entre C17 y C20 para formar acetato de testosterona. El acetato se liberaría dejando testosterona, que se oxida a AD para, finalmente, volver a ser sustrato de la BVMO y así formar testololactona (Fig. I7). Por otro lado, Abdel-Kareem et al. (2024) plantean que la BVMO de *Penicillium chrysogenum* Ras3009 introduce un 17 α -hidroxilo que permite eliminar posteriormente la cadena lateral formando testosterona. Esta testosterona se oxida a AD que, una vez más, vuelve a ser sustrato de la BVMO para formar testololactona (Fig. I7). Estas BVMOs también se han encontrado en bacterias, como es el caso de la Gram-positiva *Rhodococcus rhodochrous*, cuya estructura tridimensional y actividad enzimática se han descrito (Franceschini et al., 2012; Miyamoto et al., 1995; Morii et al., 1999). Esta monooxigenasa transforma la progesterona en acetato de testosterona, pero, a diferencia de sus homólogas fúngicas, no es capaz de lactonizar el AD. A pesar de que estos estudios confirman la actividad de algunas enzimas de varios microorganismos para utilizar la progesterona como sustrato, el catabolismo de este compuesto no está descrito.

3.1.3. Degradación aeróbica de estrógenos

La primera cepa aislada capaz de utilizar E2 como única fuente de carbono y de energía fue *Alcaligenes* sp. M21 (Payne y Talalay, 1985). Desde entonces, un número creciente de bacterias pertenecientes a los filos Proteobacteria y Actinobacteria se han ido describiendo como degradadoras de estrógenos (Chen et al., 2017; Fahrbach et al., 2006; Fujii et al., 2003; Kurisu et al., 2010; Li et al., 2020; Qin et al., 2016; Wang et al., 2019; Yu et al., 2007). En algunos de los casos estudiados, las bacterias necesitaban ser cultivadas en un medio rico para metabolizar el E2, como *Vibrio* sp. H5 (Sang et al., 2012) y *Buttiauxella* sp. S19-1 (Zhang et al., 2011), o crecer en un cultivo mixto definido de co-degradadores (e.g. *Achromobacter xylosoxidans* y *Ralstonia* sp.) (Weber et al., 2005).

Sphingomonas sp. KC8 y *C. tardaugs* NBRC 16725 son dos α -proteobacterias, ambas aisladas de fangos activos por su capacidad de degradar estrógenos, y cuyo metabolismo de la E2 se ha estudiado en profundidad (Chen et al., 2017; Fujii et al., 2003; Ibero et al., 2020; Wu et al., 2019; Yu et al., 2007). Esta ruta de degradación se denomina 4,5-*seco* y se inicia con la 17 β -deshidrogenación de la E2 para formar E1 por una 17 β -HSD (Fig. I8). A continuación, un CYP450 con actividad E1 4-hidroxilasa hidroxila el E1 y produce 4-hidroxiestrona (4-OHE1), que a su vez es sustrato de la 4-OHE1 4,5-dioxigenasa que rompe en *meta* el anillo A y lo abre. El producto de rotura es inestable y se ha visto que puede reaccionar con una molécula de amoníaco y recircularizarse de forma espontánea en ácido piridinestrónico como subproducto. En cambio, cuando el producto de rotura es sustrato de una descarboxilasa, este se activa con CoA y entra en varios pasos de hidratación, deshidenación y liasas que abren el anillo B y, finalmente, acaban en el intermediario central HIP. A partir de aquí, los anillos C y D se degradan como se ha indicado previamente (Horinouchi et al., 2005; Kieslich, 1985; van der Geize et al., 2011).

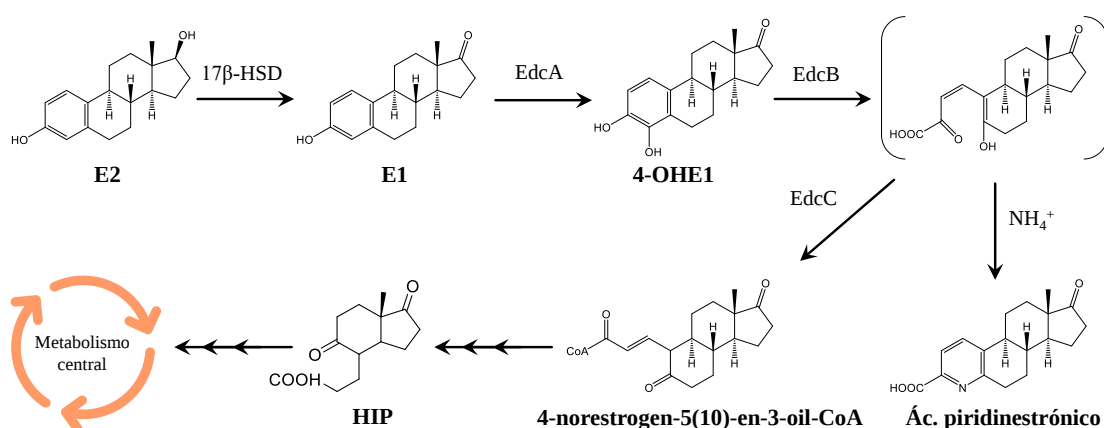


Figura I8. Ruta aerobia de degradación de estrógenos en *C. tardaugs* NBRC 16725. E2: estradiol; E1: estrona; 4-OHE1: 4-hidroxiestrona; HIP: dioxo-1,2,3,4,10,19-hexanorandrostano-5-oico; 17 β -HSD: 17 β -hidroxiesteroide deshidrogenasa; EdcA: Citocromo P450 estrona 4-hidroxilasa; EdcB: 4-hidroxiestrona 4,5-dioxigenasa; EdcC: descarboxilasa.

En *Sphingomonas* sp. KC8 se han identificado dos *clusters* (I y II) implicados en la apertura de los anillos A y B de la E2 y se ha propuesto un *cluster* III responsable de degradar los anillos C y D. Por otro lado, el *cluster* edc (del inglés *estrogen degrading cluster*) de *C. tardaogens* NBRC 16725 es similar al *cluster* II de *Sphingomonas* sp. KC8 y se compone de dos operones divergentes y un represor transcripcional TetR (Ibero et al., 2020).

4. El futuro de la biotecnología de esteroides

La carrera por desarrollar sintonas a una mayor velocidad y con mayor eficiencia, pero rebajando costes y su impacto en el medioambiente, ha llevado a la industria farmacéutica a adoptar nuevos métodos de producción como son las biotransformaciones de esteroides. A pesar de que se han alcanzado grandes avances en la biocatálisis de esteroides llegando a conseguir modificaciones regio- y estéreo-selectivas, la búsqueda de nuevas cepas o enzimas todavía continúa. El diseño racional ha permitido obtener cepas de micobacterias modificadas genéticamente que acumulan sintonas esteroideas de interés sobreexpresando o delecionando distintas enzimas involucradas en su metabolismo, como pueden ser ChoM, KstD, Ksh, FadA5, Cyp125 o Hsd4A (Zhao et al., 2021). Sin embargo, actualmente se requieren nuevas estrategias que permitan una mayor optimización de los procesos basadas en la ingeniería metabólica de sistemas, una tecnología que proporciona un análisis extensivo de todos los puntos críticos que influyen la producción de un compuesto (Zhao et al., 2021).

Uno de los problemas que podemos encontrar a la hora de aplicar chasis bacterianos es la pérdida de eficiencia catalítica cuando el proceso se escala a nivel industrial, debido a la falta de correlación entre la eficiencia obtenida en el laboratorio y en un biorreactor industrial (Calero y Nikel, 2019). Por ello, las rutas metabólicas de los esteroides, el metabolismo central, el estado redox intracelular y la estructura de la membrana celular son los aspectos fundamentales a continuar estudiando para seleccionar y perfeccionar las micobacterias como factoría celular. Al mismo tiempo, la llamada ortogonalidad metabólica es uno de los principales objetivos de la ingeniería metabólica y consiste en minimizar o evitar completamente interacciones entre la nueva ruta sintética aplicada en la factoría celular y su metabolismo basal, para así optimizar la bioproducción (Calero y Nikel, 2019). Para conseguirlo, es importante combinar las técnicas genómicas, transcriptómicas, proteómicas, metabolómicas y de modelaje computacional con las herramientas genéticas disponibles para diseñar nuevas rutas aplicables a chasis bacterianos mejorados.

Se ha abierto por tanto una nueva revolución en la industria esteroidea para alcanzar el máximo potencial que proporcionan los biocatalizadores, ya sean las factorías celulares o enzimas como los citocromos CYP450, cuya gran variedad de reacciones y amplia variabilidad de sustratos les otorga una gran versatilidad para su aplicación industrial. El objetivo final es que estos biocatalizadores acaben siendo competitivos con

la actual industria química, desarrollando bioprocesos más robustos con un mayor rendimiento y selectividad del producto.

El aumento de la producción de fármacos esteroideos por su alta demanda no puede conseguirse en detrimento del medioambiente y se requiere un compromiso para conseguir una producción más verde o evitar la contaminación de los ecosistemas. De todas las tecnologías aplicadas para la eliminación de los EDCs, los métodos biológicos constituyen una alternativa rentable, eficaz y más respetuosa con el medio ambiente, ya que evitan generar compuestos secundarios contaminantes al mineralizarlos al completo, llegándose a poner la bioaugmentación, la bioestimulación y el compostaje en el punto de mira para mejorar el tratamiento de aguas y sólidos contaminados.

Los microorganismos y su gran versatilidad fisiológica y metabólica, junto con su gran capacidad de adaptarse a la presencia de compuestos tóxicos, los convierte en los candidatos perfectos como herramientas para descontaminar residuos tóxicos. Es por ello que hay que continuar identificando bacterias degradadoras de EDCs y estudiar su relación con otros microorganismos, ya sea por sinergia o por cometabolismo, que permita potenciar su capacidad biorremediadora. Además, se deben continuar con los estudios sobre el impacto endocrino de los contaminantes para descubrir así potenciales EDCs que no estén catalogados todavía.

El compostaje, donde se encuentran comunidades microbianas que degradan eficientemente distintos tipos de contaminantes, ofrece varias ventajas para tratar grandes volúmenes de residuos y su posterior reutilización como abono, pero la eficiencia de todos los tratamientos biológicos depende de una gran variedad de factores como el tipo de residuo a descontaminar, la temperatura, el pH, la presencia de otros compuestos tóxicos, la ratio carbono:nitrógeno o la disponibilidad de oxígeno, además de requerir de grandes extensiones para realizar estos procesos (Cai et al., 2025; He et al., 2024). En este contexto, los futuros esfuerzos se deben centrar en desarrollar estrategias de remediación más efectivas que se puedan escalar para grandes volúmenes de residuos y, para ello, se necesita un mayor estudio de los mecanismos y factores que influyen en estos tratamientos, además de desarrollar nuevos materiales y tecnologías.

La biología sintética y la bioingeniería expanden las perspectivas de la biorremediación, planteándose nuevas técnicas que incluyen el diseño enzimático, la edición génica y las predicciones computacionales que nos permiten descubrir el potencial de rutas de degradación para cada EDC específico. El uso de enzimas extracelulares de origen bacteriano o fúngico proporciona varias ventajas en el tratamiento de aguas residuales, ya que permite una reacción más rápida y más sencilla de mantener y, además, no requiere manejar y operar con grandes volúmenes de lodos (Moukhtari et al., 2023). Con el desarrollo de nuevas tecnologías y/o combinando las existentes, potencialmente se podrá llegar a adoptar métodos más respetuosos con el medio ambiente y más efectivos en la eliminación de EDCs, disminuyendo su impacto nocivo en los ecosistemas.



IV. OBJETIVOS

La producción de progesterona a escala industrial mediante procesos fermentativos es uno de los desafíos biotecnológicos más relevantes para la industria de los esteroides. Su obtención actual a nivel industrial implica el empleo de un gran número de pasos de síntesis química, lo que da como resultado un proceso que, además de ser altamente contaminante, ofrece un bajo rendimiento. Por eso, en esta Tesis Doctoral se plantea el desarrollo de procesos biotecnológicos que sean menos contaminantes y más eficientes que la síntesis química actual, utilizando como producto de partida esteroides naturales y biofactorías celulares para su transformación en progesterona. En este sentido, existen algunos compuestos como el 4-HBC que actúan como sintonas para la síntesis de progesterona y que pueden obtenerse a partir de biotransformaciones de esteroides naturales por fermentación.

Para desarrollar estas biofactorías celulares es imprescindible disponer de herramientas genéticas y conocer tanto las rutas metabólicas de los esteroides precursores como las de sus intermediarios, además de las denominadas rutas laterales. En este caso es necesario estudiar el metabolismo de la progesterona para bloquear sus posibles rutas de degradación y favorecer su acumulación en las factorías celulares. Por otro lado, conocer la ruta de degradación de la progesterona es de gran valor teniendo en cuenta que este compuesto puede actuar como disruptor endocrino y, por ello, el conocimiento de la ruta podría servir para desarrollar procesos de descontaminación/biorremediación de aguas residuales.

Por último, resultados preliminares a esta Tesis Doctoral sugerían que en algunas bacterias podría existir un nuevo mecanismo para la degradación de colesterol que generase como producto intermediario progesterona, de manera similar a como se metaboliza el colesterol en humanos, lo que abriría la posibilidad de desarrollar un nuevo proceso de producción de progesterona a partir de esteroides naturales.

Es por ello que al inicio de esta Tesis Doctoral se propusieron los siguientes objetivos:

1. Producir progesterona de forma biológica a partir de esteroides naturales utilizando como biofactoría celular *M. smegmatis*. Dentro de este objetivo se han propuesto los siguientes subobjetivos:
 - a.- Estudio de las enzimas implicadas en la producción del compuesto 4-HBC a partir de esteroides naturales. Mejora de la producción de la sintona AD mediante la eliminación de la ruta de síntesis de 4-HBC.
 - b.- Producción de la sintona 4-HBC mediante la sobreexpresión de las enzimas de su ruta biosintética.
 - c.- Desarrollo de nuevas herramientas de ingeniería genética que permitan modular y controlar de manera más precisa la expresión génica en la biofactoría celular.
 - d.- Producción de progesterona mediante la expresión de sistemas de citocromos heterólogos del tipo P450sc.

2. Descubrir la ruta catabólica de la progesterona en *C. tardaugens* para usar esta bacteria en la producción de progesterona a partir de esteroides. Dentro de este objetivo se han propuesto los siguientes subobjetivos:
 - a.- Utilización de técnicas ómicas para descubrir los genes implicados en el metabolismo de la progesterona.
 - b.- Estudios genéticos y bioquímicos de los genes implicados en la ruta de degradación de la progesterona.
3. Estudiar la relación del catabolismo del colesterol de *C. tardaugens* con el metabolismo de la progesterona. Dentro de este objetivo se han propuesto los siguientes subobjetivos:
 - a.- Estudio del CYP2113 implicado en el metabolismo del colesterol.
 - b.- Diseño de mutantes que permitan la producción de progesterona a partir de esteroides.
 - c.- Utilización de técnicas ómicas para descubrir otros posibles genes implicados en el metabolismo del colesterol.



V. MATERIALES Y MÉTODOS

1. Cepas bacterianas, plásmidos, oligonucleótidos y reactivos

Las cepas bacterianas empleadas en esta Tesis Doctoral se encuentran recogidas en la Tabla M1, incluyendo sus genotipos y detalles más relevantes. Los plásmidos utilizados se detallan en la Tabla M2. Respecto a los oligonucleótidos usados en este trabajo, todos ellos fueron sintetizados por IDT (Integrated DNA Technologies, EEUU) y se encuentran en la Tabla M3, indicando su secuencia y su aplicación en el estudio. Los reactivos utilizados en este trabajo provienen de la empresa Merck (Alemania) salvo que se especifique lo contrario.

Tabla M1. Cepas bacterianas utilizadas en este trabajo.

Cepa	Genotipo y/o descripción	Referencia
<i>E. coli</i>		
HB101	Cepa huésped para el plásmido pRK600 para conjugación triparental. <i>supE44</i> , <i>ara14</i> , <i>galK2</i> , <i>leuB</i> , <i>lacY1</i> , $\Delta(gpt-proA)62$, <i>rpsL20</i> , <i>xyl-5</i> , <i>mtl-1</i> , <i>recA13</i> , $\Delta(mcrC-mrr)$, <i>hsdS20</i> , (rB ⁻ mB ⁻), SmR	(Sambrook y Russell, 2001)
DH10B	F ⁻ , <i>mcrA</i> , $\Delta(mrr\ hsdRMS-mcrBC)$, $\phi 80dlacZ\Delta M15$, $\Delta lacX74$, <i>deoR</i> , <i>recA1</i> , <i>araD139</i> , $\Delta(ara-leu)7697$, <i>galU</i> , <i>galK</i> , λ^- , <i>rpsL</i> , <i>endA1</i> , <i>nupG</i>	Invitrogen
BL21 (DE3)	F ⁻ , <i>ompT</i> , <i>hsdSB</i> (rB ⁻ mB ⁻) <i>gal</i> , <i>dcm</i> λ DE3 (con el gen 1 de la ARN polimerasa del fago T7 bajo el promotor PlacUV5)	(Sambrook y Russell, 2001)
DH5 α	Cepa huésped para clonaje. F ⁻ , λ^- , <i>endAI</i> , <i>hsdRJ7</i> (rk ⁻ , mk ⁺), <i>supE44</i> , <i>thi1</i> , <i>recA1</i> , <i>gyrA96</i> , <i>relA1</i> , <i>deoR</i> , $\Delta(lacZYA-argF)U169$, $\phi 80dlacZ\Delta M15$	(Grant et al., 1990)
W	Cepa silvestre	ATCC 9637
W (pSEVA25g19)	<i>E. coli</i> W con el plásmido pSEVA25g19	Este trabajo
W (p25_linker_RFP)	<i>E. coli</i> W con el plásmido p25_linker_RFP	Colección del laboratorio
W (p25_J23100_RFP)	<i>E. coli</i> W con el plásmido p25_J23100_RFP	Colección del laboratorio

Materiales y métodos

W (p25_J23102_RFP)	<i>E. coli</i> W con el plásmido p25_J23102_RFP	Colección del laboratorio
W (p25_J23106_RFP)	<i>E. coli</i> W con el plásmido p25_J23106_RFP	Colección del laboratorio
W (p25_Pem7_RFP)	<i>E. coli</i> W con el plásmido p25_Pem7_RFP	Colección del laboratorio
W (p25_BG17_RFP)	<i>E. coli</i> W con el plásmido p25_BG17_RFP	Colección del laboratorio
W (p25_BG37_RFP)	<i>E. coli</i> W con el plásmido p25_BG37_RFP	Colección del laboratorio
W (p25_ST_RFP)	<i>E. coli</i> W con el plásmido p25_ST_RFP	Colección del laboratorio
W (p25_B0034_RFP)	<i>E. coli</i> W con el plásmido p25_B0034_RFP	Colección del laboratorio
W (p25_BCD12_RFP)	<i>E. coli</i> W con el plásmido p25_BCD12_RFP	Colección del laboratorio
W (p25_BCD2_RFP)	<i>E. coli</i> W con el plásmido p25_BCD2_RFP	Colección del laboratorio
W (pSEVA65g19)	<i>E. coli</i> W con el plásmido pSEVA65g19	Este trabajo
W (p65_LacI_Ptrc_RFP)	<i>E. coli</i> W con el plásmido p65_LacI_Ptrc_RFP	Colección del laboratorio
W (p65_RhaSR_PrhaBAD_RFP)	<i>E. coli</i> W con el plásmido p65_RhaSR_PrhaBAD_RFP	Colección del laboratorio
W (p65_XylS_Pm_RFP)	<i>E. coli</i> W con el plásmido p65_XylS_Pm_RFP	Colección del laboratorio
W (p65_RFP_linker_GFP)	<i>E. coli</i> W con el plásmido p65_RFP_linker_GFP	Colección del laboratorio

W (p65_RFP_B0014_GFP)	<i>E. coli</i> W con el plásmido p65_RFP_B0014_GFP	Colección del laboratorio
W (p65_RFP_B1006_GFP)	<i>E. coli</i> W con el plásmido p65_RFP_B1006_GFP	Colección del laboratorio
W (p65_RFP_rpoC_GFP)	<i>E. coli</i> W con el plásmido p65_RFP_rpoC_GFP	Colección del laboratorio
W (p65_RFP_rnpB_GFP)	<i>E. coli</i> W con el plásmido p65_RFP_rnpB_GFP	Colección del laboratorio
W (p65_RFP_T500_GFP)	<i>E. coli</i> W con el plásmido p65_RFP_T500_GFP	Colección del laboratorio
<i>M. smegmatis</i>		
mc ² 155	<i>ept-1</i> , mc ² 6 mutante eficiente para electroporación	(Snapper et al., 1990)
MS6039-5941	<i>M. smegmatis</i> mc ² 155 Δ MSMEG_6039 (Δ kshB1) y Δ MSMEG_5941 (Δ kstD1)	(Galán et al., 2017)
MS6039-5941-6561	<i>M. smegmatis</i> mc ² 155 Δ MSMEG_6039 (Δ kshB1), Δ MSMEG_5941 (Δ kstD1) y Δ MSMEG_6561 (Δ msSal)	Este trabajo
MS6039-5941-6561 (pMVmsSal)	<i>M. smegmatis</i> mc ² 155 Δ MSMEG_6039 (Δ kshB1), Δ MSMEG_5941 (Δ kstD1) y Δ MSMEG_6561 (Δ msSal) con el plásmido pMV6561	Este trabajo
MS6039-5941-6561 (pMV261)	<i>M. smegmatis</i> mc ² 155 Δ MSMEG_6039 (Δ kshB1), Δ MSMEG_5941 (Δ kstD1) y Δ MSMEG_6561 (Δ msSal) con el plásmido pMV261	Este trabajo
MS6039-5941-5903	<i>M. smegmatis</i> mc ² 155 Δ MSMEG_6039 (Δ kshB1), Δ MSMEG_5941 (Δ kstD1) y Δ MSMEG_5603 (Δ hsd4A)	Este trabajo
MS6039-5941-5903 (pMV261)	<i>M. smegmatis</i> mc ² 155 Δ MSMEG_6039 (Δ kshB1), Δ MSMEG_5941 (Δ kstD1) y	Este trabajo

	Δ MSMEG_5603 (Δ hsd4A) con el plásmido pMV261	
MS6039-5941-5903 (pMVmsSal)	<i>M. smegmatis</i> mc ² 155 Δ MSMEG_6039 (Δ kshB1), Δ MSMEG_5941 (Δ kstD1) y Δ MSMEG_5603 (Δ hsd4A) con el plásmido pMVmsSal	Este trabajo
MS6039-5941-5903 (pMVmsSal-msRed)	<i>M. smegmatis</i> mc ² 155 Δ MSMEG_6039 (Δ kshB1), Δ MSMEG_5941 (Δ kstD1) y Δ MSMEG_5603 (Δ hsd4A) con el plásmido pMVmsSal-msRed	Este trabajo
MS6039-5941-5903 (pMVmsSal-msOpccR)	<i>M. smegmatis</i> mc ² 155 Δ MSMEG_6039 (Δ kshB1), Δ MSMEG_5941 (Δ kstD1) y Δ MSMEG_5603 (Δ hsd4A) con el plásmido pMVmsSal-msOpccR	Este trabajo
Δ Cyp125A3 Δ Cyp142	<i>M. smegmatis</i> mc ² 155 Δ MSMEG_5995 (Δ cyp125A3) y Δ MSMEG_5918 (Δ cyp142A2)	(García-Fernández et al., 2013)
Δ Cyp125A3- Δ 142A2- Δ 125A4	<i>M. smegmatis</i> mc ² 155 Δ MSMEG_5995 (Δ cyp125A3), Δ MSMEG_5918 (Δ cyp142A2) y Δ MSMEG_3524 (Δ cyp125A4)	Este trabajo
Δ Cyp125A3- Δ 142A2- Δ 125A4 (pSEVA65g19)	<i>M. smegmatis</i> mc ² 155 Δ MSMEG_5995 (Δ cyp125A3), Δ MSMEG_5918 (Δ cyp142A2) y Δ MSMEG_3524 (Δ cyp125A4) con el plásmido pSEVA65g19	Este trabajo
Δ Cyp125A3- Δ 142A2- Δ 125A4 (p25_Phsp60_CYP)	<i>M. smegmatis</i> mc ² 155 Δ MSMEG_5995 (Δ cyp125A3), Δ MSMEG_5918 (Δ cyp142A2) y Δ MSMEG_3524 (Δ cyp125A4) con el plásmido p25_hsp60_CYP	Este trabajo
Δ Cyp125A3- Δ 142A2- Δ 125A4 (p65_Phsp60_CYP_Fdx)	<i>M. smegmatis</i> mc ² 155 Δ MSMEG_5995 (Δ cyp125A3), Δ MSMEG_5918 (Δ cyp142A2) y Δ MSMEG_3524 (Δ cyp125A4) con el plásmido p65_hsp60_CYP_Fdx	Este trabajo
Δ Cyp125A3- Δ 142A2- Δ 125A4 (p65_Phsp60_CYP_Fdx_ArR)	<i>M. smegmatis</i> mc ² 155 Δ MSMEG_5995 (Δ cyp125A3), Δ MSMEG_5918 (Δ cyp142A2) y Δ MSMEG_3524 (Δ cyp125A4) con el plásmido p65_hsp60_CYP_Fdx_ArR	Este trabajo

<i>ΔCyp125A3-Δ142A2-Δ125A4</i> (p25_J23100_CYP)	<i>M. smegmatis</i> mc ² 155 <i>ΔMSMEG_5995</i> (<i>Δcyp125A3</i>), <i>ΔMSMEG_5918</i> (<i>Δcyp142A2</i>) y <i>ΔMSMEG_3524</i> (<i>Δcyp125A4</i>) con el plásmido p25_J23100_CYP	Este trabajo
<i>ΔCyp125A3-Δ142A2-Δ125A4</i> (p65_J23100_CYP_Fdx)	<i>M. smegmatis</i> mc ² 155 <i>ΔMSMEG_5995</i> (<i>Δcyp125A3</i>), <i>ΔMSMEG_5918</i> (<i>Δcyp142A2</i>) y <i>ΔMSMEG_3524</i> (<i>Δcyp125A4</i>) con el plásmido p65_J23100_CYP_Fdx	Este trabajo
<i>ΔCyp125A3-Δ142A2-Δ125A4</i> (p65_J23100_CYP_Fdx_ArR)	<i>M. smegmatis</i> mc ² 155 <i>ΔMSMEG_5995</i> (<i>Δcyp125A3</i>), <i>ΔMSMEG_5918</i> (<i>Δcyp142A2</i>) y <i>ΔMSMEG_3524</i> (<i>Δcyp125A4</i>) con el plásmido p65_J23100_CYP_Fdx_ArR	Este trabajo
<i>ΔCyp125A3-Δ142A2-Δ125A4</i> (p65_J23100_CYP_Phsp60_Fdx_ArR)	<i>M. smegmatis</i> mc ² 155 <i>ΔMSMEG_5995</i> (<i>Δcyp125A3</i>), <i>ΔMSMEG_5918</i> (<i>Δcyp142A2</i>) y <i>ΔMSMEG_3524</i> (<i>Δcyp125A4</i>) con el plásmido p65_J23100_CYP_Phsp60_Fdx_Ar R	Este trabajo
<i>M. smegmatis</i> (p25_linker_RFP)	<i>M. smegmatis</i> mc ² 155 con el plásmido p25_linker_RFP	Este trabajo
<i>M. smegmatis</i> (p25_Phsp60_RFP)	<i>M. smegmatis</i> mc ² 155 con el plásmido p25_Phsp60_RFP	Este trabajo
<i>M. smegmatis</i> (p25_PhspX_RFP)	<i>M. smegmatis</i> mc ² 155 con el plásmido p25_PhspX_RFP	Este trabajo
<i>M. smegmatis</i> (p25_J23100_RFP)	<i>M. smegmatis</i> mc ² 155 con el plásmido p25_J23100_RFP	Este trabajo
<i>M. smegmatis</i> (p25_J23102_RFP)	<i>M. smegmatis</i> mc ² 155 con el plásmido p25_J23102_RFP	Este trabajo
<i>M. smegmatis</i> (p25_J23106_RFP)	<i>M. smegmatis</i> mc ² 155 con el plásmido p25_J23106_RFP	Este trabajo
<i>M. smegmatis</i> (p25_Pem7_RFP)	<i>M. smegmatis</i> mc ² 155 con el plásmido p25_Pem7_RFP	Este trabajo
<i>M. smegmatis</i> (p25_Pem7_GFP)	<i>M. smegmatis</i> mc ² 155 con el plásmido p25_Pem7_GFP	Este trabajo
<i>M. smegmatis</i> (p25_BG17_RFP)	<i>M. smegmatis</i> mc ² 155 con el plásmido p25_BG17_RFP	Este trabajo
<i>M. smegmatis</i> (p25_BG37_RFP)	<i>M. smegmatis</i> mc ² 155 con el plásmido p25_BG37_RFP	Este trabajo

Materiales y métodos

<i>M. smegmatis</i> (p25_Pimyc_RFP)	<i>M. smegmatis</i> mc ² 155 con el plásmido p25_Pimyc_RFP	Este trabajo
<i>M. smegmatis</i> (p25_ST_RFP)	<i>M. smegmatis</i> mc ² 155 con el plásmido p25_ST_RFP	Este trabajo
<i>M. smegmatis</i> (p25_B0034_RFP)	<i>M. smegmatis</i> mc ² 155 con el plásmido p25_B0034_RFP	Este trabajo
<i>M. smegmatis</i> (p25_BCD12_RFP)	<i>M. smegmatis</i> mc ² 155 con el plásmido p25_BCD12_RFP	Este trabajo
<i>M. smegmatis</i> (p25_BCD2_RFP)	<i>M. smegmatis</i> mc ² 155 con el plásmido p25_BCD2_RFP	Este trabajo
<i>M. smegmatis</i> (p65_LacI_Ptrc_RFP)	<i>M. smegmatis</i> mc ² 155 con el plásmido p65_LacI_Ptrc_RFP	Este trabajo
<i>M. smegmatis</i> (p65_RhaSR_PrhaBAD_RFP)	<i>M. smegmatis</i> mc ² 155 con el plásmido p65_RhaSR_PrhaBAD_RFP	Este trabajo
<i>M. smegmatis</i> (p65_XylS_Pm_RFP)	<i>M. smegmatis</i> mc ² 155 con el plásmido p65_XylS_Pm_RFP	Este trabajo
<i>M. smegmatis</i> (p65_TetR_Ptet_RFP)	<i>M. smegmatis</i> mc ² 155 con el plásmido p65_TetR_Ptet_RFP	Este trabajo
<i>M. smegmatis</i> (p65_RFP_linker_GFP)	<i>M. smegmatis</i> mc ² 155 con el plásmido p65_RFP_linker_GFP	Este trabajo
<i>M. smegmatis</i> (p65_RFP_B0014_GFP)	<i>M. smegmatis</i> mc ² 155 con el plásmido p65_RFP_B0014_GFP	Este trabajo
<i>M. smegmatis</i> (p65_RFP_B1006_GFP)	<i>M. smegmatis</i> mc ² 155 con el plásmido p65_RFP_B1006_GFP	Este trabajo
<i>M. smegmatis</i> (p65_RFP_rpoC_GFP)	<i>M. smegmatis</i> mc ² 155 con el plásmido p65_RFP_rpoC_GFP	Este trabajo
<i>M. smegmatis</i> (p65_RFP_rnpB_GFP)	<i>M. smegmatis</i> mc ² 155 con el plásmido p65_RFP_rnpB_GFP	Este trabajo
<i>M. smegmatis</i> (p65_RFP_T500_GFP)	<i>M. smegmatis</i> mc ² 155 con el plásmido p65_RFP_T500_GFP	Este trabajo
<i>C. tarDAUGENS</i>		
NBRC 16725	Cepa silvestre	(Fujii et al., 2003)
ΔBVMO	<i>C. tarDAUGENS</i> NBRC 16725 ΔEGO55_13845	Este trabajo
ΔLLM	<i>C. tarDAUGENS</i> NBRC 16725 ΔEGO55_13860	Este trabajo

ΔLLM-BVMO	<i>C. tarдаugens</i> NBRC 16725 ΔEGO55_13845 y ΔEGO55_13860	Este trabajo
ΔLLM-BVMO (pSEVA23_BVMO-LLM)	<i>C. tarдаugens</i> NBRC 16725 ΔEGO55_13845 y ΔEGO55_13860 con el plásmido pSEVA23pLexA- EGO55_13845-13860	Este trabajo
ΔLLM-BVMO (pSEVA237plexA)	<i>C. tarдаugens</i> NBRC 16725 ΔEGO55_13845 y ΔEGO55_13860 con el plásmido pSEVA237pLexA	Este trabajo
ΔCYP2113	<i>C. tarдаugens</i> NBRC 16725 ΔEGO55_10200	(Ibero, 2022)

Tabla M2. Plásmidos utilizados en este trabajo.

Plásmidos	Descripción	Referencia
pJQ200x	Plásmido suicida utilizado para realizar mutagénesis dirigida de genes en <i>M. smegmatis</i> . Ori P15A, <i>sacB</i> , GmR, <i>xylE</i>	(Jackson et al., 2001)
pJQ3524	Plásmido pJQ200x con un fragmento aguas arriba y otro aguas abajo del gen <i>MSMEG_3524</i> clonados entre <i>XbaI</i> y <i>SacI</i>	Este trabajo
pJQ5903	Plásmido pJQ200x con un fragmento aguas arriba y otro aguas abajo del gen <i>MSMEG_5903</i> clonados entre <i>XbaI</i> y <i>SacI</i>	Este trabajo
pJQ6561	Plásmido pJQ200x con un fragmento aguas arriba y otro aguas abajo del gen <i>MSMEG_6561</i> clonados entre <i>XbaI</i> y <i>SacI</i>	Este trabajo
pMV261	Plásmido lanzadera <i>Mycobacterium/E. coli</i> con KmR y el promotor Phsp60 de <i>M. tuberculosis</i>	(Stover et al., 1991)
pMVmsSal	Plásmido pMV261 con el gen <i>MSMEG_6561</i> que codifica msSal de <i>M. smegmatis</i> y clonado entre <i>BamHI</i> - <i>HindIII</i>	(Hernández-Fernández et al., 2023)
pMVmsSal-msRed	Plásmido pMV261 con los genes <i>MSMEG_6561</i> y <i>MSMEG_6563</i> que codifican msSal y msRed de <i>M. smegmatis</i> , respectivamente, y clonado entre <i>BamHI</i> - <i>HindIII</i> - <i>HpaI</i>	Este trabajo

pMVmsSal-msOpccR	Plásmido pMV261 con los genes <i>MSMEG_6561</i> y <i>MSMEG_1623</i> que codifican msSal y msOpccR de <i>M. smegmatis</i> , respectivamente, y clonado entre <i>Bam</i> HI- <i>Hind</i> III- <i>Hpa</i> I	Este trabajo
pRK600	Plásmido necesario para las conjugaciones triparentales; CmR, ColE1 <i>ori</i> V, RP4 <i>ori</i> T	(Sharma y Signer, 1990)
pK18msg	Plásmido suicida para realizar mutagénesis dirigida por recombinación homóloga adaptado de pK18mobsacB para clonar con <i>Bsa</i> I; KmR, ColE <i>ori</i> V, Mob+, <i>lacZ</i> α , <i>sacB</i>	(Godoy et al., 2023)
pK18EGO55_13845	Plásmido pK18msg con un fragmento aguas arriba y otro aguas abajo del gen <i>EGO55_13845</i> clonado con <i>Bsa</i> I	Este trabajo
pK18EGO55_13860	Plásmido pK18msg con un fragmento aguas arriba y otro aguas abajo del gen <i>EGO55_13860</i> clonado con <i>Bsa</i> I	Este trabajo
pSEVA237plexA	Plásmido que expresa de forma constitutiva el gen <i>gfp</i> bajo el promotor PlexA; KmR, pBBR1 <i>ori</i> V	(Ibero et al., 2021)
pSEVA23_BVMO-LLM	Plásmido que expresa de forma constitutiva los genes <i>EGO55_13845</i> y <i>EGO55_13860</i> bajo el promotor PlexA y clonados entre <i>Xba</i> I- <i>Sph</i> I y <i>Hind</i> III- <i>Spe</i> I	Este trabajo
pET-29a(+)	Plásmido de expresión; KmR, ColE1 <i>ori</i> V, promotor T7	Novagen
pET-BVMO	Plásmido pET-29 con el gen <i>EGO55_13845</i> clonado entre <i>Nde</i> I- <i>Xho</i> I	Este trabajo
pET-LLM	Plásmido pET-29 con el gen <i>EGO55_13860</i> clonado entre <i>Nde</i> I- <i>Xho</i> I	Este trabajo
pET-CYP2113	Plásmido pET-29 con el gen <i>EGO55_10200</i>	Este trabajo
pET-CYP2113_Fdx_ArR	Plásmido pET-29 con los genes <i>EGO55_10200</i> , <i>Saro_3659</i> y <i>Saro_0216</i>	Este trabajo

pL0_AB	Plásmido vacío de nivel 0 con los sitios de fusión A y B para clonar por GS; SmR	(Blázquez et al., 2023)
pL0_DG	Plásmido vacío de nivel 0 con los sitios de fusión D y G para clonar por GS; SmR	(Blázquez et al., 2023)
pL0_Pimyc_AB	Plásmido de nivel 0 con el promotor Pimyc entre los sitios de fusión A y B para clonar por GS; SmR	Este trabajo
pL0_PhspX_AB	Plásmido de nivel 0 con el promotor PhspX entre los sitios de fusión A y B para clonar por GS; SmR	Este trabajo
pL0_Phsp60_AB	Plásmido de nivel 0 con el promotor Phsp60 entre los sitios de fusión A y B para clonar por GS; SmR	Este trabajo
pL0_J23100_AB	Plásmido de nivel 0 con el promotor J23100 entre los sitios de fusión A y B para clonar por GS; SmR	(Blázquez et al., 2023)
pL0_RBS-Std_BD	Plásmido de nivel 0 con el RBS estándar entre los sitios de fusión B y D para clonar por GS; SmR	(Blázquez et al., 2023)
pL0_CYP_DG	Plásmido de nivel 0 con el gen <i>Saro_3659</i> entre los sitios de fusión D y G para clonar por GS; SmR	Este trabajo
pL0_Fdx_DG	Plásmido de nivel 0 con el gen <i>Saro_3658</i> entre los sitios de fusión D y G para clonar por GS; SmR	Este trabajo
pL0_ArR_DG	Plásmido de nivel 0 con el gen <i>Saro_0216</i> entre los sitios de fusión D y G para clonar por GS; SmR	Este trabajo
pL0_Bba_B1006_GI	Plásmido de nivel 0 con el terminador sintético B1006 entre los sitios de fusión G e I para clonar por GS; SmR	(Blázquez et al., 2023)
pSEVA25g19	Plásmido vacío de nivel 1 sin dianas para <i>BsaI</i> y <i>BbsI</i> para clonar por GS; KmR, RSF1010 <i>ori</i>	Este trabajo
p25_linker_RFP	Plásmido de nivel 1 con un linker, el RBS estándar, la <i>Red Fluorecent Protein</i> y el terminador B1006	Colección del laboratorio
p25_Phsp60_RFP	Plásmido de nivel 1 con el promotor Phsp60, el RBS estándar, la <i>Red</i>	Este trabajo

	<i>Fluorescent Protein</i> y el terminador B1006	
p25_PhspX_RFP	Plásmido de nivel 1 con el promotor PhspX, el RBS estándar, la <i>Red Fluorescent Protein</i> y el terminador B1006	Este trabajo
p25_J23100_RFP	Plásmido de nivel 1 con el promotor J23100, el RBS estándar, la <i>Red Fluorescent Protein</i> y el terminador B1006	Colección del laboratorio
p25_J23102_RFP	Plásmido de nivel 1 con el promotor J23102, el RBS estándar, la <i>Red Fluorescent Protein</i> y el terminador B1006	Colección del laboratorio
p25_J23106_RFP	Plásmido de nivel 1 con el promotor J23106, el RBS estándar, la <i>Red Fluorescent Protein</i> y el terminador B1006	Colección del laboratorio
p25_Pem7_RFP	Plásmido de nivel 1 con el promotor Pem7, el RBS estándar, la <i>Red Fluorescent Protein</i> y el terminador B1006	Colección del laboratorio
p25_BG17_RFP	Plásmido de nivel 1 con el promotor BG17, el RBS estándar, la <i>Red Fluorescent Protein</i> y el terminador B1006	Colección del laboratorio
p25_BG37_RFP	Plásmido de nivel 1 con el promotor BG37, el RBS estándar, la <i>Red Fluorescent Protein</i> y el terminador B1006	Colección del laboratorio
p25_Pimyc_RFP	Plásmido de nivel 1 con el promotor Pimyc, el RBS estándar, la <i>Red Fluorescent Protein</i> y el terminador B1006	Este trabajo
p25_Phsp60_CYP	Plásmido de nivel 1 con el promotor Phsp60, el RBS estándar, el gen <i>Saro_3659</i> y el terminador B1006	Este trabajo
p25_Phsp60_Fdx	Plásmido de nivel 1 con el promotor Phsp60, el RBS estándar, el gen <i>Saro_3658</i> y el terminador B1006	Este trabajo
p25_Phsp60_ArR	Plásmido de nivel 1 con el promotor Phsp60, el RBS estándar, el gen <i>Saro_0216</i> y el terminador B1006	Este trabajo

p25_J23100_CYP	Plásmido de nivel 1 con el promotor J23100, el RBS estándar, el gen <i>Saro_3659</i> y el terminador B1006	Este trabajo
p25_J23100_Fdx	Plásmido de nivel 1 con el promotor J23100, el RBS estándar, el gen <i>Saro_3658</i> y el terminador B1006	Este trabajo
p25_J23100_ArR	Plásmido de nivel 1 con el promotor J23100, el RBS estándar, el gen <i>Saro_0216</i> y el terminador B1006	Este trabajo
p25_ST_RFP	Plásmido de nivel 1 con el promotor Pem7, el RBS estándar, la <i>Red Fluorescent Protein</i> y el terminador B1006	Colección del laboratorio
p25_B0034_RFP	Plásmido de nivel 1 con el promotor Pem7, el RBS B0034, la <i>Red Fluorescent Protein</i> y el terminador B1006	Colección del laboratorio
p25_BCD12_RFP	Plásmido de nivel 1 con el promotor Pem7, el RBS BCD12, la <i>Red Fluorescent Protein</i> y el terminador B1006	Colección del laboratorio
p25_BCD2_RFP	Plásmido de nivel 1 con el promotor Pem7, el RBS BCD2, la <i>Red Fluorescent Protein</i> y el terminador B1006	Colección del laboratorio
pSEVA65g19	Plásmido vacío de nivel 2 sin dianas para <i>BsaI</i> y <i>BbsI</i> para clonar por GS; GmR, RSF1010 <i>ori</i>	Este trabajo
p65_Phsp60_CYP_Fdx	Plásmido de nivel 2 con los genes <i>Saro_3658</i> y <i>Saro_3659</i> en unidades transcripcionales independientes bajo los promotores Phsp60, RBS estándar y terminadores B1006	Este trabajo
p65_Phsp60_CYP_Fdx_ArR	Plásmido de nivel 2 con los genes <i>Saro_3658</i> , <i>Saro_3659</i> y <i>Saro_0216</i> en unidades transcripcionales independientes bajo los promotores Phsp60, RBS estándar y terminadores B1006	Este trabajo
p65_J23100_CYP_Fdx	Plásmido de nivel 2 con los genes <i>Saro_3658</i> y <i>Saro_3659</i> en unidades transcripcionales independientes bajo los promotores J23100, RBS estándar y terminadores B1006	Este trabajo

p65_J23100_CYP_Fdx_ArR	Plásmido de nivel 2 con los genes <i>Saro_3658</i> , <i>Saro_3659</i> y <i>Saro_0216</i> en unidades transcripcionales independientes bajo los promotores J23100, RBS estándar y terminadores B1006	Este trabajo
p65_J23100_CYP_Phsp60_Fdx_ArR	Plásmido de nivel 2 con el gen <i>Saro_3658</i> bajo el promotor J23100, y los genes <i>Saro_3659</i> y <i>Saro_0216</i> en unidades transcripcionales independientes bajo el promotor Phsp60. Todos con el RBS estándar y terminadores B1006	Este trabajo
p65_LacI_Ptrc_RFP	Plásmido de nivel 2 con el gen <i>lacI</i> y la <i>Red Fluorescent Protein</i> bajo el promotor Ptrc. Inducible por IPTG	Colección del laboratorio
p65_RhaSR_PrhaBAD_RFP	Plásmido de nivel 2 con el gen <i>RhaSR</i> y la <i>Red Fluorescent Protein</i> bajo el promotor PrhaBAD. Inducible por ramnosa	Colección del laboratorio
p65_XylS_Pm_RFP	Plásmido de nivel 2 con el gen <i>XylS</i> y la <i>Red Fluorescent Protein</i> bajo el promotor Pm. Inducible por ácido <i>m</i> -toluico	Colección del laboratorio
p65_TetR_Ptet_RFP	Plásmido de nivel 2 con el gen <i>TetR</i> y la <i>Red Fluorescent Protein</i> bajo el promotor Ptet. Inducible por anhidrotetraciclina	Este trabajo
p65_RFP_linker_GFP	Plásmido de nivel 2 con el promotor Pem7, la <i>Red Fluorescent Protein</i> , un linker y la <i>Green Fluorescent Protein</i>	Colección del laboratorio
p65_RFP_B0014_GFP	Plásmido de nivel 2 con el promotor Pem7, la <i>Red Fluorescent Protein</i> , el terminador B0014 y la <i>Green Fluorescent Protein</i>	Colección del laboratorio
p65_RFP_B1006_GFP	Plásmido de nivel 2 con el promotor Pem7, la <i>Red Fluorescent Protein</i> , el terminador B1006 y la <i>Green Fluorescent Protein</i>	Colección del laboratorio
p65_RFP_rpoC_GFP	Plásmido de nivel 2 con el promotor Pem7, la <i>Red Fluorescent Protein</i> , el terminador rpoC y la <i>Green Fluorescent Protein</i>	Colección del laboratorio

p65_RFP_rnpB_GFP	Plásmido de nivel 2 con el promotor Pem7, la <i>Red Fluorescent Protein</i> , el terminador rnpB y la <i>Green Fluorescent Protein</i>	Colección del laboratorio
p65_RFP_T500_GFP	Plásmido de nivel 2 con el promotor Pem7, la <i>Red Fluorescent Protein</i> , el terminador T500 y la <i>Green Fluorescent Protein</i>	Colección del laboratorio

Tabla M3. Oligonucleótidos empleados en este trabajo.

Oligonucleótido	Descripción	Secuencia (5'-3')
MSMEG_6561_UP_Fw	Amplificación del fragmento aguas arriba de <i>MSMEG_6561</i> (<i>SacI</i> - <i>Bam</i> HI)	CGGGGAGCTCGTCGGGACGG
MSMEG_6561_UP_Rv		TCGCAGAACACG
MSMEG_6561_DW_Fw	Amplificación del fragmento aguas abajo de <i>MSMEG_6561</i> (<i>Bam</i> HI- <i>Xba</i> I)	CCCTGGATCCCACGACCTGAC
MSMEG_6561_DW_Rv		CTGACCTTCCGCGGC
MSMEG_3524_UP_Fw	Amplificación del fragmento aguas arriba de <i>MSMEG_3524</i> (<i>SacI</i> - <i>Bam</i> HI)	ACTTGGATCCATTTGCCCGCA
MSMEG_3524_UP_Rv		CGCCGAGGT
MSMEG_3524_DW_Fw	Amplificación del fragmento aguas abajo de <i>MSMEG_3524</i> (<i>Bam</i> HI- <i>Xba</i> I)	GGCTTCTAGAACATCGACCGC
MSMEG_3524_DW_Rv		ATGGGTGCCCT
MSMEG_3524_ext_Fw	Amplificación de la región mutada <i>MSMEG_3524</i> para comprobación	GTGAGAGCTCCGGATCGGCC
MSMEG_3524_ext_Rv		AGCGC
MSMEG_5903_UP_Fw	Amplificación del fragmento aguas arriba de <i>MSMEG_5903</i> (<i>Xho</i> I- <i>Avr</i> II)	TTAAGAATTCCTTCGGGATAA
MSMEG_5903_UP_Rv		CGCGGCAGC
MSMEG_5903_DW_Fw	Amplificación del fragmento aguas abajo de <i>MSMEG_5903</i> (<i>Xho</i> I- <i>Avr</i> II)	ATCGGAATTCACACGTCGGGT
MSMEG_5903_DW_Rv		TCGGCGGCA
MSMEG_5903_ext_Fw	Amplificación de la región mutada <i>MSMEG_5903</i> para comprobación	ACTCTCTAGATCTCACCGATG
MSMEG_5903_ext_Rv		GCGTGCCCGTA
MSMEG_5903_DW_Fw	Amplificación del fragmento aguas abajo de <i>MSMEG_5903</i> (<i>Xho</i> I- <i>Avr</i> II)	TCACGCAGATACGACATGGT
MSMEG_5903_DW_Rv		AACAGGTTCCACTCGCGATC
MSMEG_5903_DW_Fw	Amplificación del fragmento aguas abajo de <i>MSMEG_5903</i> (<i>Xho</i> I- <i>Avr</i> II)	5CTAGCTCGAGCTCACGGATC
MSMEG_5903_DW_Rv		GTGCGCGCTCC
MSMEG_5903_DW_Fw	Amplificación del fragmento aguas abajo de <i>MSMEG_5903</i> (<i>Xho</i> I- <i>Avr</i> II)	CTAGCCTAGGCGGCGACTGCA
MSMEG_5903_DW_Rv		CTGATGGAG
MSMEG_5903_DW_Fw	Amplificación del fragmento aguas abajo de <i>MSMEG_5903</i> (<i>Xho</i> I- <i>Avr</i> II)	CTAGCCTAGGCGGACAGGTC
MSMEG_5903_DW_Rv		GAGATCGGC

MSMEG_5903_DW_Rv	MSMEG_5903 (AvrII- <i>Xho</i> I)	CTAGCTCGAGCATCTGCGCGA CGGTCTCGC
MSMEG_6563_Fw	Amplificación del gen <i>MSMEG_6563</i> incluyendo un RBS (<i>Hind</i> III- <i>Hpa</i> I)	ACCAAGCTTGACCTAAGGAG GTGAATATGACAGCGACGA
MSMEG_6563_Rv		5TAGGATGTAACTCATGTAC GCCTCCTCGACGC
MSMEG_1623_Fw	Amplificación del gen <i>MSMEG_1623</i> incluyendo un RBS (<i>Hind</i> III- <i>Hpa</i> I)	CTCAAGCTTGACCTAAGGAGG TGAATATGCGCTATGTCGTT
MSMEG_1623_Rv		CGCTGTAAACACGACTACCAG TGCACGCCGG
EGO55_13845UpFw	Amplificación del fragmento aguas arriba del gen <i>EGO55_13845</i> (<i>Bsa</i> I)	CACATGGTCTCTTCTTGTCGA TATCACCGATCA
EGO55_13845UpRv		CACATGGTCTCGTCATCCCCG TTACTGTTCTGAT
EGO55_13845DwFw	Amplificación del fragmento aguas abajo del gen <i>EGO55_13845</i> (<i>Bsa</i> I)	CACATGGTCTCGATGACTTGT CTGATATCTGTTGGA
EGO55_13845DwRv		CACATGGTCTCGTTCCCACTT ATGACACGAAAGC
EGO55_13860UpFw	Amplificación del fragmento aguas arriba del gen <i>EGO55_13860</i> (<i>Bsa</i> I)	CACATGGTCTCTTCTTGGTCG GCAATGATC
EGO55_13860UpRv		ACATGGTCTCGTCATCATGCC CTGATTTATGTC
EGO55_13860DwFw	Amplificación del fragmento aguas abajo del gen <i>EGO55_13860</i> (<i>Bsa</i> I)	CACATGGTCTCGATGACCTTC TCTCCCGAGCT
EGO55_13860DwRv		CACATGGTCTCGTTCCCGAAC AACTCTGCCTAT
EGO55_13845_ext_Fw	Amplificación de la región mutada <i>EGO55_13845</i> para comprobación	AAGTGCGTTATCGTTATC
EGO55_13845_ext_Rv		AATTGGTCAGCATTCTC
EGO55_13860_ext_Fw	Amplificación de la región mutada <i>EGO55_13860</i> para comprobación	TGAGGCGATCAGATC
EGO55_13860_ext_Rv		CTACGCCGAATATGTC
EGO55_13845FwXbaI	Amplificación de <i>EGO55_13845</i> (<i>Xba</i> I- <i>Sph</i> I)	ATATTCTAGAAAGGAGGTGA ATATGTCCTCTGTGAATAACG GGT
EGO55_13845RvSphI		ATATGCATGCCTAGCTGATGA CGTAATCGG

EGO55_13860FwHindIII	Amplificación de <i>EGO55_13860</i> (<i>HindIII</i> - <i>SpeI</i>)	ATATAAGCTTAAGGAGGTGA ATATGAAGTTTTCCATCATTT ATGAA
EGO55_13860RvSpeI		ATATACTAGTTCAGCCAGCCA CAGTTTCAA

2. Medios y condiciones de cultivo

Las soluciones y medios de cultivo líquidos y sólidos empleados se prepararon con agua destilada y se esterilizaron por calor húmedo en un autoclave a 121 °C y 1 atm de presión o por filtración con filtros estériles Millipore® de 0.22 µm. Todos los crecimientos se monitorizaron por turbidimetría a DO₆₀₀ con un espectrofotómetro UVMini-1240 (Shimadzu, Japón).

2.1. Medios y condiciones de cultivo para *Escherichia coli*

Para cultivar las células de *E. coli* se utilizó el medio líquido Luria-Bertani (LB) (Sambrook y Russell, 2001) a 37 °C y en un agitador orbital a 200 rpm. Como medio sólido se utilizaron placas de LB con agar 1,5% (p/v) de Bacto Agar (Pronadisa, España). Cuando fue necesario, se usaron antibióticos a distintas concentraciones: kanamicina a 50 µg/mL, gentamicina a 10 µg/mL y cloranfenicol a 20 µg/mL (Sigma, EEUU).

2.2. Medios y condiciones de cultivo para *Caenibius tardaugens*

El cultivo de *C. tardaugens* NBRC16725 en medio rico se realizó en 16 g/L de *Nutrient Broth* (NB) (Difco, EEUU) en un agitador orbital a 200 rpm a 30 °C. Como medio mínimo para ensayar distintas fuentes de carbono, se utilizó el medio M63 (por litro: 13,6 g KH₂PO₄, 20 g (NH₄)₂SO₄, 0,5 mg FeSO₄·7H₂O, 0,25 g MgSO₄·7H₂O; pH 8) complementado con 0,04 g/L de CaCl₂. Como medio sólido se utilizaron placas de NB con agar 1,5% (p/v) de Bacto Agar (Pronadisa, España). Cuando fue necesario, se añadió kanamicina a 10 µg/mL, gentamicina a 10 µg/mL o rifampicina a 50 µg/mL (Sigma, EEUU).

2.3. Medios y condiciones de cultivo para *Mycolicibacterium smegmatis*

El cultivo de *M. smegmatis* en medio rico se realizó en LB suplementado con Tween 80 (Sigma) al 0,05% (v/v) para evitar la agregación de las células. Como medio mínimo se adaptó el medio *Middlebrook 7H9 Broth* (Difco, EEUU). Para ello, al preparado comercial se le añadió la fuente de carbono deseada en vez de ADC (albúmina, dextrosa y catalasa), como indica la casa comercial, y se suplementó con Tween 80 al

0,05%. *M. smegmatis* se cultivó a 37 °C y a 200 rpm en un agitador orbital. El medio sólido empleado fue *Bacto Middlebrook 7H10 Agar* (Difco, EEUU) para cultivar las células y, en caso necesario, se usaron los siguientes antibióticos: kanamicina 20 µg/mL y gentamicina 5 µg/mL.

El medio de cultivo empleado en los biorreactores para *M. smegmatis* se compone de licor de maíz macerado (Sigma, EEUU) (20 g/L), NaNO₃ (5,4 g/L), NH₄Cl (0,28 g/L) y NaH₂PO₄ (0,28 g/L) disuelto en agua destilada ajustado a pH 7,4.

2.4. Preparación de esteroides

Los esteroides se pueden disolver utilizando metil-β-ciclodextrina (CDX) (TRMB-T *Randomly Methylated-Beta-Cyclodextrine*, CTD Inc., EEUU). La CDX se preparó a 70 mM en tampón PBS (por litro: 8 g NaCl, 0,2 g KCl, 1,44 g Na₂HPO₄ y 0,24 g KH₂PO₄; pH 6,8) y a continuación se añadieron los esteroides a las siguientes concentraciones: testosterona (TES) a 10 mM, progesterona (PROG) a 9 mM y colesterol (CHO) y colesteno (CHONA) a 7 mM, de forma que sus concentraciones finales fueron 2 mM TES, 1,71 mM PROG, 1,33 mM CHO y 1,33 mM CHONA. La solución final se sonicó durante 30 min y después se filtró como se ha descrito previamente. Con esta proporción, la concentración final de las CDX en las soluciones stock fue de 13,33 mM y la concentración de los esteroides fue equimolar en cantidad de carbono.

En otros casos los esteroides se disolvieron en Tyloxapol (Sigma, EEUU) 10% (v/v) a una concentración final de 5 mM. Una vez preparada, la solución se sonicó y se autoclavó 2 veces para su completa disolución.

Los fitoesteres de soja usados como sustrato en los ensayos con biorreactores se obtuvieron de Gadea Biopharma (España) y se componen principalmente de β-sitosterol (54,4%), campesterol (36,8%) y estigmasterol (8,7%). Para su disolución, los fitoesteres se disolvieron en aceite de girasol sonicando y calentando a 50 °C durante una hora y a continuación la mezcla se autoclavó.

2.5. Conservación de las cepas bacterianas

Las cepas bacterianas se conservaron en sus respectivos medios sólidos a 4 °C durante periodos de un mes máximo. Para su conservación a largo plazo, todas las cepas se crecieron en medio rico líquido al que se añadió glicerol 20% (v/v) y posteriormente se guardaron a -80 °C.

3. Técnicas de manipulación de ADN

En las técnicas de biología molecular para manipulación del ADN se utilizaron endonucleasas de restricción, enzima T4 ADN ligasa y polimerasa *Phusion High-Fidelity*

de New England Biolabs (EEUU), siguiendo las especificaciones marcadas por la casa comercial. Para purificar fragmentos de ADN, se emplearon los kits *QIAquick PCR Purification Kit* (Qiagen, Alemania) o *QIAquick Gel Extraction Kit* (Qiagen, Alemania). En caso de medir la concentración de ADN, se utilizó un *NanoDrop™ OneC* (Thermo Scientific, EEUU) de acuerdo a las instrucciones del equipo.

3.1. Extracción de ADN cromosómico

Para la extracción de ADN cromosómico se utilizó el siguiente procedimiento. Un volumen de 10 mL de medio inoculado y cultivado durante 18 h se centrifugó a 3800 rpm en una centrífuga 5810 R (Eppendorf, EEUU) durante 15 min. El sedimento se resuspendió en 400 µL de Tris-EDTA (TE) (10 mM Tris pH 7,5 con 1 mM EDTA) y se incubó a 80 °C durante 20 min. Una vez atemperado, se añadieron 50 µL de lisozima (Sigma, EEUU) a 10 mg/mL, se mezcló con un vórtex y se dejó incubar durante 1 h a 37 °C. A continuación, se añadieron 75 µL de proteinasa K (Roche, Suiza) a 10 mg/mL con 10% (v/v) de dodecilsulfato sódico (SDS), se mezcló con vórtex y se incubó durante toda la noche a 65 °C. A la mezcla se le añadieron 200 µL de NaCl 5M precalentado a 65 °C y se volvió a incubar a la misma temperatura 10 min. Una vez pasado el tiempo, se añadieron 750 µL de cloroformo:alcohol isoamílico (24:1), se homogeneizó mediante vórtex y se centrifugó en una MiniSpin (Eppendorf, EEUU) a 13400 rpm durante 10 min. El ADN, que se encuentra en la fase acuosa, se transfirió a otro tubo eppendorf y se precipitó con 0,6 volúmenes de isopropanol frío. Se mezcló invirtiendo el tubo y se dejó 30 min a -20 °C. Se volvió a centrifugar 15 min a 13400 rpm, se retiró el sobrenadante y se lavó el sedimento con etanol al 70% (v/v) en frío. Se dejó secar 5 min a temperatura ambiente y finalmente se resuspendió en 40-100 µL de agua MilliQ.

3.2. Aislamiento de ADN plasmídico

Para la extracción de ADN plásmidico de *E. coli* o de *C. tardegens* se utilizó el kit *High Pure Plasmid Purification Kit* (Roche, Suiza) siguiendo las especificaciones del fabricante. En el caso de *M. smegmatis*, antes de utilizar el kit indicado, se incubaron las células previamente con 50 µL de lisozima 50 mg/mL a 37 °C durante 30 min.

3.3. Reacción en cadena de la polimerasa de ADN (PCR)

Para la amplificación de ADN se utilizó la polimerasa *Phusion High-Fidelity* (New England Biolabs, EEUU) incluyendo su tampón de reacción. Además, se añadieron oligonucleótidos 0,5 µM y dNTPs 0,25 mM y DMSO al 5% (v/v) cuando se usa un ADN como molde con un alto contenido en GC. Las condiciones estándar de PCR fueron: un primer ciclo de 98 °C durante 30 s; 30 ciclos de 10 s a 98 °C, 30 s a la temperatura de anillamiento de los oligonucleótidos seguido de una amplificación a 72 °C (dejando 30 s por cada 1 kb); un ciclo final de 5 min a 72 °C. Todo esto se llevó a cabo en el

termociclador *Veriti™ 96-Well* (Applied Biosystems, EEUU). Como molde se utilizó ADN genómico o plasmídico extraído o biomasa de la cepa de interés hervida en TE a 80 °C durante 20 min.

3.4. Electroforesis de ADN en gel de agarosa

Los fragmentos de ADN extraídos o purificados se visualizaron en geles de agarosa (Condalab, España) al 0,7% o 1,5% (p/v) en tampón TAE (Tris-HCl 40 mM, ácido acético 20 mM y EDTA 2 mM; pH 8.1). Este mismo tampón fue utilizado como electrolito en una cubeta Mupid-One a 135 V durante 25 min. A las muestras se les añadió el tampón de carga *10X Loading Buffer* (Takara) y los geles se tiñeron con *Gel Red Nucleic Acid Gel Stain* (Biotium, EEUU). El marcador de peso molecular utilizado fue *NZYDNA Ladder III* (NzyTech, Portugal) y los geles se revelaron con radiación UV en un transiluminador Gel Doc XR+ (Bio-Rad, EEUU).

3.5. Secuenciación de ADN

La secuencia de fragmentos de ADN se llevó a cabo por la empresa Plasmidsaurus Inc. (EEUU) mediante tecnología de nanoporos (Oxford Nanopore Technologies, Reino Unido) o por la empresa SECUGEN SL con un secuenciador automático ABI PRISM 3730 (Applied Biosystems, EEUU). Las reacciones de amplificación se realizaron mediante PCR en el termociclador Gene Amp PCR System 2400 (Perkin-Elmer, EEUU) y para la reacción de secuenciación se usó el kit *Dye Terminator Cycle Sequencing Ready Reaction Kit* (Applied Biosystems, EEUU) con la polimerasa AmpliTaq FS (Thermo Scientific, EEUU).

3.6. Clonaje *Golden Standard*

Para el clonaje modular se ha utilizado la herramienta *Golden Standard* descrita por (Blázquez et al., 2023). Las distintas partes a clonar correspondientes a promotor, RBS, CDS, terminador o *linker* que contienen las dianas de *BbsI* se clonan de forma individual en un plásmido de nivel 0 y se seleccionan por resistencia a espectinomicina. Para formar una unidad transcripcional, estas partes se clonan a la vez en un plásmido de nivel 1, que presenta resistencia a kanamicina. Para clonar varias unidades transcripcionales construidas en plásmidos de nivel 1, se usa un plásmido de nivel 2 y se seleccionan por resistencia a gentamicina. La reacción se explica brevemente: para 20 µL de mezcla de reacción se añaden 40 fmol de cada plásmido y 200 U de ligasa T4 (New England Biolabs, EEUU). Para construir los plásmidos de nivel 0 y 2 se emplean 10 U de la enzima de restricción *BbsI* (New England Biolabs, EEUU), mientras que para ensamblar los plásmidos de nivel 1 se utilizan 10 U de *BsaI*-HFv2 (New England Biolabs, EEUU). Todas las reacciones se realizan con buffer CutSmart 1X (New England Biolabs, EEUU) y con ATP 10 µM. Las condiciones de reacción son: 28 ciclos de 37 °C durante

1,5 min más 3 min a 16 °C. Después 5 min a 50 °C, 5 min a 37 °C y finalmente 10 min a 80 °C. Finalmente todo el volumen se usa para transformar *E. coli* como se describe a continuación.

4. Experimentos de modificación genética

4.1. Transformación de células de *E. coli*

Las células de *E. coli* se transformaron mediante dos métodos. El primero de ellos es el método de RbCl descrito por (Sambrook y Russell, 2001) con choque térmico, incubando primero 3 min a 37 °C, seguido de una segunda incubación en hielo durante 2 min y una posterior recuperación con LB a 37 °C y agitación durante 1 h. El segundo método es la electroporación. En este caso las células a transformar se cultivaron hasta la fase estacionaria y después se lavaron 3 veces con agua destilada fría para retirar las sales presentes en el medio. A continuación, las células se concentraron en 100 µL y se colocaron en una cubeta de electroporación de 2 mm (Cell Projects, Reino Unido), previamente enfriada a 4 °C, junto con 200 ng de ADN a transformar. La mezcla se incubó en frío durante 10 min y, una vez pasado el tiempo, se electroporó usando un *Gene Pulser* (Bio-Rad, EEUU) a 2,5 kV, 25 µF y 200 Ω. Inmediatamente se añadió 1 mL de medio LB a la cubeta y se incubó 1 h a 37 °C en agitación. Finalmente, las células transformadas se plaquearon en el medio sólido indicado con los antibióticos correspondientes.

4.2. Transformación de células de *C. tardaugens*

Para modificar genéticamente la cepa *C. tardaugens*, se utilizó la conjugación triparental. En este caso, *C. tardaugens* resistente a rifampicina fue la cepa receptora, la cepa *E. coli* DH10B con el plásmido a transferir fue la cepa donadora y la cepa *E. coli* HB101 con el plásmido pRK600 fue la cepa auxiliar (de Lorenzo y Timmis, 1994). La cepa receptora primero se cultivó en medio NB hasta su fase exponencial tardía (48 h aprox.). A continuación, las células se centrifugaron a 3800 rpm a 4 °C durante 15 min utilizando una centrífuga 5810 R (Eppendorf, EEUU) y el sedimento se lavó con solución salina estéril 0,85% (p/v) de NaCl y se volvieron a centrifugar. El sedimento final se resuspendió en 800 µL de esta misma solución. Por otro lado, la cepa donadora y la cepa auxiliar se cultivaron en 2 mL de LB en tubos estériles de plástico de 10 mL hasta la fase exponencial tardía (6 h aprox.) usando un preinóculo cultivado en 10 mL de LB a 37 °C y 200 rpm con los antibióticos necesarios. Los cultivos se lavaron como se ha descrito previamente y se resuspendieron en 100 µL de solución salina. Se mezclaron 50 µL de cada cepa y se colocaron en un filtro de disco MF-Millipore de 0,22 µm (Merck, Alemania) sobre una placa de NB, que se llevó a 30 °C durante un día. Finalmente, se recogieron los filtros, se colocaron en un tubo eppendorf con 1 mL de solución salina y se agitó en un vórtex para separar las células, que fueron seleccionadas plaqueando en placas de NB con rifampicina 50 µg/mL y el antibiótico correspondiente.

4.3. Transformación de células de *M. smegmatis*

Las células de *M. smegmatis* se transformaron mediante electroporación. Para preparar las células electrocompetentes, se siguió el protocolo descrito por Parish y Stoker, (1998). Con un preinóculo de la cepa en fase estacionaria en LB con 0,05% de Tween 80 se inocularon 200 mL del mismo medio de cultivo a una DO_{600} de 0,01, y se deja crecer a 37 °C a 200 rpm de agitación durante unas 18 h ($DO_{600}=0,8-1$). Este cultivo se incubó durante 1 h a 4 °C y se centrifugó a 3800 rpm durante 20 min en una centrífuga 5810 R (Eppendorf, EEUU) para después realizar varios lavados y centrifugaciones seguidas usando glicerol 10% (v/v) con Tween 80 0,05%. Tras el último lavado y centrifugación, el sedimento se resuspendió en 1 mL de la misma solución de lavado y se repartió en alícuotas de 200 μ L que se guardaron a -80 °C hasta su uso.

Una vez generadas las células electrocompetentes, 100 μ L de estas se mezclaron con 1 μ g del ADN plasmídico a transformar y se incubaron en una cubeta de electroporación de 2 mm (Cell Projects, Reino Unido) en hielo durante 10 min. Después, las células se electroporaron en un *Gene Pulser* (Bio-Rad, EEUU) a 2,5 kV, 25 μ F y 1000 Ω y se devolvieron al hielo durante 5 min más. Finalmente, las células se resuspendieron en 1 mL de LB con Tween 80 0,05% y se incubaron a 37 °C en agitación durante 4 h hasta que se sembraron en una placa de LB con el antibiótico correspondiente.

4.4. Obtención de mutantes por delección génica

La construcción de las diferentes cepas mutantes recogidas en la Tabla M1 se realizó mediante doble recombinación homóloga usando los plásmidos suicidas pJQ200x y pK18msg para *M. smegmatis* y *C. tar daugens*, respectivamente. Para ello se amplificaron por PCR dos fragmentos de unos 500-800 pb correspondientes a la región aguas arriba (UP) y aguas abajo (DOWN) del gen a mutar usando los oligonucleótidos recogidos en la Tabla M3. Ambos fragmentos UP y DOWN y los vectores suicidas de destino se digirieron usando las enzimas de restricción correspondientes siguiendo las instrucciones del fabricante y posteriormente se ligaron utilizando la ligasa del fago T4 (New England Biolabs, EEUU). Esta ligación se transformó en *E. coli* como se ha indicado previamente y se seleccionó e identificó mediante PCR la cepa con el plásmido deseado.

En *C. tar daugens*, una vez obtenido el plásmido suicida con los fragmentos UP y DOWN necesarios para mutar un gen, se realizó una conjugación triparental y los transformantes que han realizado la primera recombinación homóloga se seleccionaron en placas de NB con kanamicina (100 μ g/mL) y rifampicina (50 μ g/mL). La inserción del plásmido se comprobó mediante PCR usando un oligonucleótido que hibrida en el plásmido (R24 o F24) y otro que hibrida en la región externa de la inserción (Tabla M3). El transformante seleccionado se cultivó en medio líquido NB sin antibióticos hasta la fase estacionaria y a continuación se plaqueó en placas de NB con sacarosa 5% (p/v), de

forma que la expresión del gen *sacB* del plásmido suicida hace tóxica la sacarosa y se fuerza la segunda recombinación homóloga. Las colonias obtenidas se analizaron por PCR para comprobar que tenían el nuevo genotipo mutado o que habían recuperado el genotipo silvestre utilizando dos oligonucleótidos externos al gen a mutar.

En el caso de *M. smegmatis*, se siguió un procedimiento parecido salvo que en este caso las células se electroporaron para ser transformadas. Además, los transformantes se seleccionaron en placas de LB con gentamicina (5 µg/mL) y la contraselección se realizó en medio 7H10 con sacarosa al 10%.

5. Técnicas de manipulación de ARN

5.1. Extracción de ARN total de *C. tardaugens*

La extracción de ARN se realizó a partir de 12 mL de cultivo en dos condiciones distintas en fase exponencial ($DO_{600}=0,6-0,8$). Estas células primero se lavaron con solución salina 0,85% y se pretrataron primero con 400 µL de lisozima (50 mg/mL) en tampón TE a 37 °C durante 10 min. A continuación, se siguieron 3 ciclos de congelación-descongelación incubando 4 min en hielo seco y luego 3 min a 37 °C. Una vez lisadas las células, el ARN total se aisló utilizando el kit *High Pure RNA Isolation kit* (Roche, Suiza) y para eliminar el ADN contaminante se utilizó el kit *DNase and Removal treatment kit* (Ambion, EEUU), siguiendo con ambos las indicaciones del fabricante. Para medir la concentración de ARN obtenido se utilizó un *NanoDrop™ OneC* (Thermo Scientific, EEUU).

5.2. Secuenciación masiva de ARN (RNA-seq)

La secuenciación completa del ARN se realizó en MacroGen NGS (Corea). Primero se seleccionaron aquellas muestras que presentaban un valor de *RNA Integrity Number* (RIN) igual o superior a 7 y para eliminar el ARN ribosomal se utilizó el kit de Illumina Inc. *Ribo-Zero rRNA Removal kit* (EEUU). Se preparó a continuación una librería del ARNm de 100 pb, pareadas, con el kit *TruSeq RNA Sample Prep kit v2* (Illumina Inc., EEUU) y la calidad de estas librerías se comprobó en un bioanalizador *Agilent Technologies 2100* (EEUU). Finalmente, la secuenciación de la librería se realizó en un HiSeq 3000/4000 (Illumina Inc. EEUU) con el kit *TruSeq 3000 4000 SBS kit v3* (Illumina Inc, EEUU) y los resultados obtenidos se realizaron en el servicio de Bioinformática y Bioestadística del Centro de Investigaciones Biológicas “Margarita Salas” (CIB-CSIC). La calidad de los datos crudos obtenidos primero se comprobó usando FastQC y después las lecturas se limpiaron y filtraron usando el programa Trimmomatic 0.39 (Bolger et al., 2014). Las lecturas obtenidas se alinearon frente al genoma de *C. tardaugens* (número de acceso CP034179) mediante el programa Bowtie2 2.4.2 (Anders et al., 2015) y el análisis de los genes expresados diferencialmente se

realizó con el programa Deseq2 1.32.0 en R 3.6.3 (R: The R Project for Statistical Computing; <https://www.r-project.org>).

6. Técnicas de manipulación de proteínas

6.1. Obtención de extractos proteicos de *E. coli*

Para sobreexpresar proteínas en *E. coli*, las células se cultivaron en 50 mL de LB con el antibiótico necesario en agitación a 200 rpm y 37 °C hasta alcanzar una DO₆₀₀ de 0,6-0,8 para después inducir la expresión de la proteína de interés añadiendo 0,2 mM de isopropil-β-D-tiogalactopiranosido (IPTG) (Sigma, EEUU). Una vez inducido, el cultivo se mantuvo en agitación a 200 rpm a 30 °C durante 3 h más, momento en el que se centrifugó y se lavó el cultivo y el sedimento se guardó a -20 °C hasta su posterior uso.

Para sobreexpresar el citocromo P450, a los cultivos de LB se les añadió 1 mM de ácido 5-aminolevulínico y 1 mM de FeSO₄ para favorecer la formación del grupo hemo.

Para lisar las células, el sedimento se resuspendió primero en un tampón de fosfato potásico 50 mM pH 8 y se sonicaron en un sonicador Branson 150 alternando 6 ciclos de 10 s a máxima intensidad y 30 s de reposo en hielo. También las muestras se lisaron usando una prensa de French (Aminco Corp, EEUU) a una presión de 20000 psi. En ambos casos, el extracto total se centrifugó a 14000 rpm a 4 °C durante 15 min en una centrífuga Sigma 1-15K (EEUU) para separar la fracción soluble de la fracción insoluble. La concentración total de proteína presente en la fracción soluble se determinó mediante el método de Bradford (Bradford, 1976).

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

Las muestras de proteínas primero se hirvieron durante 10 min a 100 °C en presencia de tampón ruptura (Tris-HCl 250 mM a pH 6,8, SDS 2% (p/v), β-mercaptoetanol 5% (v/v), glicerol 10% (v/v) y azul de bromofenol 0,05% (v/v)). Se preparó un gel de poliacrilamida al 12,5% (v/v) con SDS (Laemmli, 1970) y la electroforesis se realizó en una cubeta de electroforesis (Bio-Rad, EEUU) usando como electrolito tampón Tris-HCl 25 mM (pH 8) con glicina 192 mM y 0,1% (p/v) de SDS. Primero se aplicó una corriente de 80 V durante 20 min y después se subió a 120 V. Después, las proteínas fueron teñidas con *Coomasie Brilliant Blue R250* (Swank y Munkres, 1971). Para conocer el peso molecular de las proteínas, se utilizó el marcador *BlueStar Prestained Protein Marker* (Nippon Genetics, Japón) que presenta el siguiente patrón de proteínas: 180 kDa (azul), 130 kDa (azul), 100 kDa (azul), 75 kDa (rojo), 63 kDa (azul), 48 kDa (azul), 35 kDa (azul), 28 kDa (verde), 17 kDa (azul), 10 kDa (azul).

6.3. Ensayos enzimáticos *in vitro* de la actividad Baeyer-Villiger monooxigenasa (BVMO)

La actividad de la proteína BVMO se ensayó en un espectrofotómetro UV-1900i (Shimadzu, Japón) midiendo el descenso de la absorbancia del cofactor NADPH a 340 nm ($\epsilon_{\text{NADPH}} = 6,22 \text{ cm}^{-1} \text{ mM}^{-1}$) en presencia del sustrato. La reacción enzimática se realizó a 25 °C en cubetas de cuarzo de HELLMA Analytics (Alemania) de 10 mm de paso óptico añadiendo 2 mg/mL de extracto proteico en tampón fosfato 50 mM pH 8. Los sustratos a ensayar se prepararon a 50 mM en metanol y se añadieron a una concentración final de 0,5 mM y la reacción empezó cuando 0,2 mM de NADPH se añadieron al medio de reacción. La actividad específica (U/mg de proteína) se definió como el número de nmol de NADPH consumidos en 1 min por mg de proteína total.

Para medir la actividad de la proteína LLM con actividad Baeyer-Villiger monooxigenasa, se realizó un ensayo ligeramente diferente. En este caso, 4 mg/mL de extracto proteico en tampón fosfato 50 mM pH 7,5 se ensayaron en un vial de vidrio (Supelco, USA) con agitación magnética y a temperatura ambiente. Se añadió FMN a una concentración de 20 μM y el sustrato a 0,5 mM. La reacción comenzó al añadir 5 mM de NADH y se tomaron muestras de 200 μL a los 0, 5 y 30 min y la reacción se detuvo con 500 μL de cloroformo. En este caso, la actividad específica (U/mg proteína) se definió como el número de nmol de sustrato consumido en 1 min por mg de proteína total.

6.4. Ensayos enzimáticos *in vitro* de la actividad de citocromos CYP450

Para ensayar la actividad enzimática de los citocromos CYP450 se utilizaron viales de vidrio (Supelco, USA). El volumen final de reacción era de 200 μL que contenían 100 μg de extracto proteico, tampón fosfato sódico 50 mM (pH 7,5), MgCl_2 10 mM, CDX 0,4 mM, ferredoxina de espinaca 1 μM , ferredoxina NADP-reductasa (0,2 U/mg), catalasa de hígado bovino (0,1 mg/mL), glucosa 6-fosfato deshidrogenasa (0,4 U/mg) y glucosa 6-fosfato 4 mM. El sustrato se disolvió en metanol a una concentración de 5 mM y se añadió a una concentración final de 0,1 mM. La reacción se inició nada más añadir NADPH 0,1 mM, se mantuvo con agitación magnética y a temperatura ambiente y se tomaron muestras de 50 μL a los 0, 30 y 60 min. La reacción se detuvo añadiendo 100 μL de acetonitrilo con ácido fórmico 0,1% (v/v).

6.5. Modelado y análisis *in silico* de la estructura de los CYP450 candidatos de *C. tardaugens*

La estructura tridimensional de los CYP450 seleccionados se modeló mediante reemplazo molecular utilizando los modelos de AlphaFold2 y mediante varios ciclos repetitivos de refinamiento. La conformación del grupo hemo junto con el colesterol se obtuvo de la base de datos *Protein Data Bank* (PDB, <https://www.rcsb.org/>) y se añadió

de forma manual a la proteína revisando las interacciones químicas posibles. Las estructuras se visualizaron con el programa PyMOL (<https://www.pymol.org/>).

7. Biotransformaciones de esteroides *in vivo*

7.1. Ensayos de biotransformación de esteroides con células en crecimiento en matraz

Para el ensayo se partió de un preinoculo de *M. smegmatis* cultivado en LB con Tween 80 0,05% durante un fin de semana. La biomasa se lavó con solución salina con Tween 80 0,05% y se empleó para inocular el medio de biotransformación a una $DO_{600}=0,05$. Como medio líquido para las biotransformaciones se utilizó el medio mínimo *Middlebrook 7H9 Broth* (Difco) con glicerol 9 mM como fuente de carbono y energía, y el esteroide a biotransformar a una concentración de 1 mM disuelto en Tyloxapol 10%. En este caso, el Tyloxapol sustituye al Tween 80 para evitar la formación de grumos y su concentración final no superó en ningún caso el 3,6%. Se tomaron 800 μ L de muestra cada 24 h durante 96 h que se usaron para medir la DO_{600} del cultivo y se guardaron a -20 °C hasta su uso.

En el caso de las cepas de *C. tardaugens* a ensayar, las biotransformaciones se realizaron en medio líquido NB complementado con el esteroide de interés disuelto en CDX como se ha descrito en el apartado 2.4. De igual forma, se extrajeron muestras de 800 μ L para controlar el crecimiento del cultivo y analizarlas posteriormente.

7.2. Ensayos de biotransformación de esteroides con células en reposo (*resting cells*) en matraz

Para las biotransformaciones en *resting cells* con *M. smegmatis* se empleó un cultivo en medio rico para inocular a una $DO_{600}=0,05$ 50 mL de LB con Tween 80 0,05%. Este cultivo se mantuvo 48 h hasta alcanzar la fase estacionaria ($DO_{600}=4-5$), se centrifugó a 3800 rpm durante 20 min y se lavaron las células con solución salina 0,85%. El medio de biotransformación utilizado fue un tampón fosfato potásico 50 mM pH 8 con glicerol 2% (p/v) o glucosa 2% (p/v) como fuente de energía y el esteroide de interés disuelto en CDX a una concentración final de 500 mg/L. El ensayo de *resting cells* se inoculó a una DO_{600} inicial de 16-18 y se tomaron 800 μ L de muestra a las 0, 1, 3, 5, 7, 24, 30 y 48 h, que se guardaron a -20 °C hasta su análisis. Cuando se emplearon fitosteroides en el ensayo, debido a su alta hidrofobicidad, primero se disolvieron en 100 μ L de isopropanol caliente a 60 °C y después se añadió al medio con la misma concentración final de CDX.

C. tardaugens se cultivó en 125 mL de medio rico NB durante 48 h hasta una $DO_{600}=2$. Estas células se lavaron con solución salina y se concentraron mediante centrifugación a 3800 rpm en una centrífuga 5810 R (Eppendorf, EEUU) durante 15 min

hasta una DO_{600} inicial de 16-18 en medio NB complementado con 500 mg/L del sustrato esteroideo preparado en CDX. Se tomaron 800 μ L de muestra a las 0, 1, 3, 5, 7, 24 y 30 h, que se guardaron a -20 °C hasta su uso.

Los medios ricos utilizados en los inóculos se complementaron con ácido 5-aminolevulínico 1 mM y $FeSO_4$ 1 mM si la cepa de interés expresaba un CYP450.

En el caso de *E. coli*, se utilizó un preinóculo cultivado en LB con el antibiótico correspondiente. Estas células se lavaron y se usaron para inocular a una $DO_{600}=0,2$ 50 mL de medio mínimo M9 con casaminoácidos 1% (p/v), ácido 5-aminolevulínico 0,5 mM, $FeSO_4$ 0,1 mM y glucosa 0,4% (p/v). Pasadas 2 h, cuando el cultivo alcanzó una $DO_{600}=0,6-0,8$, éste se indujo con 0,2 mM de IPTG. Unas 18 h más tarde se concentraron 10 mL del cultivo en 1 mL de tampón compuesto de tampón fosfato 50 mM pH 7,5, glicerol 2% (p/v), IPTG 0,1 mM, CDX 20 mM, el sustrato a ensayar a 0,25 mM disuelto en metanol y el antibiótico correspondiente. La mezcla de reacción se incubó a 30 °C con agitación a 200 rpm y se tomaron muestras a las 0 y a las 24 h.

7.3. Ensayos de biotransformación de esteroides en crecimiento a escala de biorreactor

El medio de biotransformación en biorreactor se preparó como se indica en el apartado 2.3 y se añadió a un biorreactor de 2 L Sartorius AG (Alemania). Al mismo tiempo, se disolvieron 14 g de fitoesteroides en 160 mL de aceite de girasol como se describe en el apartado 2.4 y se autoclavaron junto al reactor. Recién autoclavados se mezclaron los fitoesteroides en aceite con el medio de cultivo y se ajustó el pH inicial a 7,4, la temperatura a 37 °C, la agitación a 600 rpm y el flujo de aire a 1 vvm. El oxígeno disuelto se mantuvo controlado por encima del 20% a lo largo del experimento. El preinóculo se cultivó en 200 mL de LB con Tween 80 0,05% durante 72 h y luego se concentró en 100 mL a una DO_{600} de 2-2,5 mediante centrifugación a 3800 rpm en una centrífuga 5810 R (Eppendorf, EEUU) durante 10 min, de forma que, al inocular, el volumen final en el biorreactor fue de 1 L y la DO_{600} inicial de 0,2-0,25. Finalmente, se fueron tomando muestras de 15 mL de cultivo cada 24 h y se guardaron a -20 °C.

8. Técnicas cromatográficas

8.1. Extracción de esteroides

Los esteroides presentes en las muestras obtenidas se extrajeron añadiendo dos volúmenes de cloroformo, la mezcla se llevó al vórtex durante 45 s y se centrifugó durante 2 min a 13000 rpm en una centrífuga MiniSpin (Eppendorf, EEUU). La fase orgánica se evaporó en una campana de extracción y el sedimento se resuspendió en 200 μ L de acetonitrilo (ACN).

En el caso de las muestras obtenidas del ensayo en biorreactor, estas se centrifugaron a 3800 rpm durante 30 min en una centrífuga 5810 R (Eppendorf, EEUU) a 3800 rpm de forma que se separan la fracción de aceite de la fracción acuosa. 50 µL de la fracción de aceite se utilizaron para su análisis posterior.

8.2. Cromatografía líquida de alta eficiencia (HPLC) y acoplada a espectrometría de masas (HPLC-MS)

La cuantificación de los fitoesteroides se realizó mediante HPLC según el protocolo descrito en Yang et al. (2023). 50 µL de muestra se disolvieron en 200 µL acetato de etilo de los cuales 20 µL se analizaron en un HPLC 1260 *Infinity II LC System* (Agilent Technologies, EEUU) acoplado a un detector de diodos (*Diode Array Detector*, DAD). La columna empleada fue la *Eclipse Plus C18* (5 µm; 4.6 × 250 mm) (Agilent Technologies, EEUU) calentada a 30 °C, la fase móvil fue metanol:acetonitrilo (80:20) y el flujo se fijó en 1 mL/min. Cada carrera duró 30 min y se detectaron los compuestos midiendo su absorbancia a 195, 205 y 210 nm.

Para analizar el resto de esteroides, las muestras se inyectaron en un HPLC-MS 1260 *Infinity II* (Agilent Technologies, EEUU) acoplado a un *InfinityLab LC/MSD XT* (Agilent Technologies, EEUU). El equipo presenta un inyector automático con DAD y una trampa iónica con electrospray. La columna utilizada fue Poroshell 120 EC-C18 (4 µm; 4.6 × 100 mm) (Agilent, EEUU). El método de separación consistió en un flujo de 1 mL/min empezando con un 70% de solvente A (H₂O con 0,1% de ácido fórmico) y un 30% de solvente B (ACN con 0,1% de ácido fórmico) durante 5 min. Después, el solvente B se incrementó hasta el 55% durante 20 min y después se volvió al 30% de solvente B en 2 min. La detección de los compuestos se realizó a 195, 210, 230, 250 y 275 nm.

Los compuestos esteroideos se identificaron utilizando los compuestos comerciales disponibles como referencia o mediante su espectro de masas. La cuantificación de los esteroides se realizó mediante rectas patrón de los compuestos puros.

8.3. Cromatografía de gases acoplada a espectrometría de masas (GC-MS)

Previo al análisis de las muestras por GC-MS, la fase orgánica extraída y evaporada se derivatizó formando derivados de éter de trimetilsililo reaccionando 50 µL de N,O-bis(trimetilsilil)trifluoroacetamida con trimetilclorosilano (BSTFA) (Sigma, EEUU) con 50 µL de piridina a 80 °C durante 20 min. Los compuestos comerciales estándar se derivatizaron de la misma manera. El equipo utilizado fue un cromatógrafo de gases Agilent 7890A acoplado al detector de masas Agilent 5975C (Agilent Technologies, EEUU). Los espectros de masa se registraron en modo de impacto de electrones (EI) a 70 eV y el rango fijado fue m/z 50-550. La columna capilar del cromatógrafo presentaba unas dimensiones de 30 x 0,25 mm (espesor de película de 0,25

µm) HP-5MS (5% de difenilo y 95% de dimetilpolisiloxano) (Agilent Technologies, EEUU). Las condiciones de trabajo fueron las siguientes: relación de división 20:1, temperatura de inyector 320 °C, la temperatura de la columna empezó en 240 °C durante 3 min y luego se calentó hasta los 320 °C incrementando 5 °C por minuto. Los espectros de masas obtenidos se compararon con aquellos obtenidos de los estándares depositados en la base de datos de espectros de masas NIST (NIST 2011).

9. Ensayos de fluorescencia en placas multipocillo

Para estudiar la actividad de los distintos promotores, sitios de unión del ribosoma (RBS) y terminadores en *M. smegmatis*, la cepa se transformó con distintas construcciones que contenían fusiones transcripcionales con la proteína roja fluorescente (*Red Fluorescent Protein*, RFP) como marcador (Tabla M2). Las cepas recombinantes se cultivaron primero en LB con Tween 80 0,05% con el correspondiente antibiótico y se utilizaron como preinóculo de cultivos en medio mínimo 7H9 con glicerol 18 mM y Tween 80 0,5% hasta una $DO_{600}=0,6-0,8$. Estos cultivos se utilizaron a su vez para preparar una placa de 96 pocillos de fondo transparente-negro (Thermo Fisher) con 3 réplicas técnicas de 200 µL a una $DO_{600}=0,2$. El crecimiento y la fluorescencia emitida se monitorizaron en un lector de placas Victor Nivo PerkinElmer (Biolab, EEUU) a 37 °C con agitación a 1200 rpm, midiendo DO_{600} y excitando a 580 nm y recogiendo a 625 nm para la RFP, con 500 ms de exposición. Las medidas se tomaron cada 30 min.

10. Cálculos

El rendimiento molar (MY) se calculó utilizando la siguiente fórmula [Eq. 1]:

$$MY (\%) = \frac{[Sintona](mM)}{[Sc](mM)} \cdot 100$$

Donde el término “Sintona” se refiere a cualquier esteroide obtenido como producto y “Sc” a los esteroides consumidos.

La Productividad (P) se calculó mediante la fórmula indicada [Eq. 2]:

$$P \left(\frac{g}{L \cdot h} \right) = \frac{[Sintona](g/L)}{t(h)}$$

Donde el término “t” se refiere a tiempo (h).

11. Análisis bioinformáticos

El software Geneious Prime® (v2025.1.3, <https://www.geneious.com>) se utilizó para todos los análisis *in silico* relativos a la visualización de genomas y plásmidos y a la comprobación de clonajes, secuenciaciones y amplificaciones por PCR. Los

alineamientos de secuencia se realizaron utilizando las herramientas BLAST (del inglés Basic local alignment search tool, <https://blast.ncbi.nlm.nih.gov/Blast.cgi>) y ClustalW (<https://www.ebi.ac.uk/jdispatcher/msa/clustalo>) aplicando los parámetros preestablecidos. Para realizar los análisis comparativos de *clusters* génicos se empleó la herramienta *online* CAGECAT (<https://cagecat.bioinformatics.nl/>) con los parámetros preestablecidos. Las representaciones gráficas de los datos y el cálculo de su media y desviación estándar se realizaron con el *software* GraphPad Prism (v 9.4.1, <https://www.graphpad.com>).



VI. RESULTADOS

CAPÍTULO 1. Producción de esteroides con interés industrial en *Mycolicibacterium smegmatis* a partir de esteroides

1.1. Optimización de la producción de androstenediona (C19) mediante biotransformación a partir de esteroides

Publicación 1 (Anexo I)

Título: *Identification of the aldolase responsible for the production of 22-hydroxy-23,24-bisnorcholesterol-4-ene-3-one from natural sterols in Mycolicibacterium smegmatis*

Autores: Gabriel Hernández-Fernández, Miguel G. Acedos, José L. García, Beatriz Galán

Revista: *Microbial Biotechnology*

Referencia: Hernández-Fernández, G., Acedos, M.G., García, J.L. & Galán, B. (2023) Identification of the aldolase responsible for the production of 22-hydroxy-23,24-bisnorcholesterol-4-ene-3-one from natural sterols in *Mycolicibacterium smegmatis*. *Microbial Biotechnology*, 00, 1–10. <https://doi.org/10.1111/1751-7915.14270>

Aportación del doctorando: el doctorando Gabriel Hernández ha llevado a cabo la parte experimental y de investigación, así como el análisis de los datos y la escritura del borrador original del artículo.

Resumen:

Las sintonas C19, como la androstenediona (AD) son de gran importancia para la industria farmacéutica ya que se utilizan como materia prima para la síntesis de testosterona y otros andrógenos y estrógenos. Actualmente se producen utilizando cepas mutantes de micobacterias a partir de la biotransformación de esteroides. Estas cepas además de producir AD acumulan otros subproductos, como los esteroides C22, entre los que se encuentra el 22-hidroxi-23,24-bisnorcholesterol-4-en-3-ona (4-HBC). Estos compuestos intermediarios se generan a través de una ruta alternativa de la degradación de los esteroides generando una disminución en el rendimiento del proceso, lo que dificulta los pasos de purificación posteriores del compuesto de interés. En este trabajo se ha identificado el gen *MSMEG_6561* que codifica una aldolasa involucrada en la transformación del 22-hidroxi-3-oxo-colesterol-4-en-24-carboxil-CoA (22-OH-BCN-CoA) en (20S)-3-oxopregn-4-en-20-carboxaldehído (3-OPA), el precursor del esteroide C22 4-HBC. Cuando se deleta este gen en una cepa de *M. smegmatis* productora de AD, la producción de la sintona C19 de interés a partir de esteroides aumenta, evitándose la acumulación de cualquier subproducto. Esta nueva cepa produce hasta 240 mg/L de AD a partir de colesterol mediante biotransformaciones en matraz, aumentando el rendimiento molar de un 59% a un 81%. Además, esta cepa se ensayó a escala de biorreactor utilizando 10 g/L de fitoesteroides como materia prima, observándose una producción de 7 g/L de AD con una productividad a las 24 h de 0,17 g/L h⁻¹.

1.2. Producción del esteroide C22 22-hidroxi-23,24-bisnorcol-4-en-3-ona mediante biotransformación de esteroides

Publicación 2 (Anexo II)

Título: *Improving the production of 22-hydroxy-23,24-bisnorchol-4-ene-3-one in Mycolicibacterium smegmatis*

Autores: Gabriel Hernández-Fernández, Miguel G. Acedos, Isabel de la Torre, Juan Ibero, José L. García, Beatriz Galán

Revista: *Microbial Biotechnology*

Referencia: Hernández-Fernández, G., Acedos, M.G., de la Torre, I., Ibero, J., García, J.L. & Galán, B. (2024) Improving the production of 22-hydroxy-23,24-bisnorchol-4-ene-3-one in *Mycolicibacterium smegmatis*. *Microbial Biotechnology*, 17, e14551. <https://doi.org/10.1111/1751-7915.14551>

Aportación del doctorando: el doctorando Gabriel Hernández ha participado en el desarrollo experimental y de investigación del trabajo, el análisis de los datos y la escritura del borrador original del artículo.

Resumen:

El compuesto 22-hidroxi-23,24-bisnorcol-4-en-3-ona (4-HBC) es una sintona C22 de gran interés farmacéutico para la producción de esteroides tales como la progesterona. Este esteroide se puede obtener como un intermediario del catabolismo de esteroides naturales, como el colesterol o los fitoesteroides, utilizando cepas mutantes de micobacterias. En este trabajo se ha obtenido un mutante de *M. smegmatis* capaz de producir 4-HBC delecionando 3 genes: *MSMEG_5903* (*Δhsd4A*), *MSMEG_6039* (*ΔkshB1*) y *MSMEG_5941* (*ΔkstD1*), que codifican una β-hidroxiacil-CoA deshidrogenasa, el componente reductasa de la cetoesteroide-9α-hidroxilasa y la 3-cetoesteroide-Δ¹-deshidrogenasa, respectivamente. Este triple mutante fue capaz de producir 5,6 g/L de la sintona C22 a partir de 14 g/L de fitoesteroides. Para mejorar este rendimiento, se ha estudiado el papel de la enzima aldehído deshidrogenasa msRed (*MSMEG_6563*) de *M. smegmatis* en el metabolismo de 4-HBC. Este gen se encuentra en un *cluster* de genes junto con la aldolasa msSal (*MSMEG_6561*), cuyo papel en el metabolismo del 4-HBC, descrito previamente, es el de catalizar la eliminación retroaldólica del acetil-CoA del 22-hidroxi-3-oxo-colest-4-en-24-carboxil-CoA (22-OH-BCN-CoA) y producir (20S)-3-oxopregn-4-en-20-carboxaldehído (3-OPA). Para demostrar la función de msRed, ésta se expresó junto con la aldolasa msSal en el plásmido de expresión pMV261, con el que se transformó el mutante triple de *M. smegmatis*. La cepa resultante mejoró la producción de 4-HBC a partir de colesterol respecto a la cepa parental, lo que sugiere que msRed es la enzima responsable de la reducción del 3-OPA a 4-HBC. Se estudió también el papel de otra reductasa descrita en *M. neoaurum*, denominada msOpccR (*MSMEG_1623*). Se observó que esta enzima es capaz de realizar la reducción del 3-OPA al 4-HBC, pero de una forma menos eficiente que msRed. Cuando

la cepa triple mutante sobreexpresa las enzimas msSal y msRed, la producción de 4-HBC aumenta hasta 12,7 g/L, mostrando una productividad de 0,185 g/L h⁻¹.

1.3. Desarrollo de nuevas herramientas genéticas en micobacterias: expansión del *Golden Standard*

En la última década el desarrollo de nuevas herramientas moleculares que permiten optimizar el uso de la ingeniería genética en los microorganismos ha aumentado considerablemente, especialmente para poder estandarizar y facilitar la construcción de circuitos génicos que ofrecen soluciones para modificar y mejorar distintos procesos celulares. Aplicaciones como el llamado *Modular Cloning* (MoClo) o los sistemas JUMP, *Multi-Kingdom* o *Cyano Gate* son de uso habitual y constituyen un gran avance para la ingeniería genética, tanto de organismos procariotas como eucariotas (Vasudevan et al., 2019; Weber et al., 2011). Estos sistemas se basan principalmente en el uso de enzimas de restricción de tipo IIS que permiten ensamblar elementos básicos modulares como promotores, sitios de unión al ribosoma (RBS), secuencias codificantes (CDS) y terminadores en una única unidad transcripcional. Las enzimas de restricción de tipo IIS se caracterizan por reconocer secuencias asimétricas de ADN y cortar fuera de la diana (Szybalski et al., 1991). Esto genera extremos cohesivos específicos de cada corte, denominados sitios de fusión (FS), y la pérdida de la diana de restricción, permitiendo realizar la digestión y la ligación de fragmentos en un mismo tubo de reacción. De esta forma los módulos básicos, también denominados partes de nivel 0, se pueden ensamblar en un determinado orden y dirección para formar una unidad transcripcional, considerada una construcción de nivel 1 (Engler et al., 2014; Weber et al., 2011).

Sin embargo, estas herramientas presentan varias limitaciones. La primera de ellas es que en ocasiones las partes de nivel 0 no son intercambiables entre sistemas, ya sea por el uso de distintas enzimas de restricción de tipo IIS o por el diseño de los FS. Además, estas herramientas solo tienen aplicación en un grupo reducido de organismos, llegando a ser incluso específicos del clado o del filo (Blázquez et al., 2023; Vasudevan et al., 2019). Es por ello que Blázquez et al. (2023) decidieron desarrollar el sistema *Golden Standard* (GS). Esta herramienta surge al combinar la tecnología modular del MoClo con la estandarización de los plásmidos pSEVA, lo que permite aumentar la portabilidad y la versatilidad del sistema. Este sistema también se basa en un ensamblaje primero de los niveles 0 en niveles 1 gracias a FS específicos que van desde la A hasta la I (Fig. R1.1). Los niveles 1 a continuación se pueden ensamblar de forma jerárquica en circuitos genéticos de niveles 2 y después de nivel 3. Estas combinaciones se consiguen alternando las dianas de *BbsI* y *BsaI* en las distintas construcciones, además de variar la resistencia del plásmido receptor entre kanamicina (niveles 1 y 3) y gentamicina (nivel 2) para facilitar la selección de la construcción. Blázquez et al. (2023) construyeron así una colección de plásmidos para GS utilizando distintos orígenes de replicación (RK2, pBBR1 y pUC) y se ensayó su portabilidad en α -, β -, γ - y δ -proteobacterias, demostrando la aplicabilidad de esta herramienta en bacterias no modelo. Sin embargo, esta

herramienta no es válida para aquellos microorganismos que no son capaces de replicar con los orígenes de replicación mencionados, como es el caso de las micobacterias.

Actualmente existen otras herramientas genéticas aplicables a micobacterias basadas en el clonaje con enzimas de restricción de tipo IIS, como el denominado clonaje por intercambio de fragmentos (Arnold et al., 2018), o basadas en el uso de partes modulares o *BioBricks* (Oliveira et al., 2019). Estas alternativas, sin embargo, son muy limitadas ya que la enzima de restricción de tipo IIS utilizada (*SapI*) deja extremos cohesivos de 3 nucleótidos, mientras que *BbsI* y *BsaI* dejan extremos de 4 nucleótidos, lo que impide la compatibilidad de partes entre otros sistemas (Arnold et al., 2018). Esto implica que, por ejemplo, los 8 promotores descritos para micobacterias por Oliveira et al. (2019) no puedan ser ensayados con otros sistemas en los que hay una colección mayor de promotores potencialmente útiles en micobacterias. Además, los plásmidos diseñados para clonar los *BioBricks* replican solo en *E. coli* y en micobacterias (Oliveira et al., 2019). Por esta razón, aprovechando la modularidad de los plásmidos pSEVA, se planteó la necesidad de expandir el sistema GS e implementar esta tecnología en Micobacterias donde las herramientas genéticas son escasas, facilitando así la ingeniería de otras bacterias no modelo.

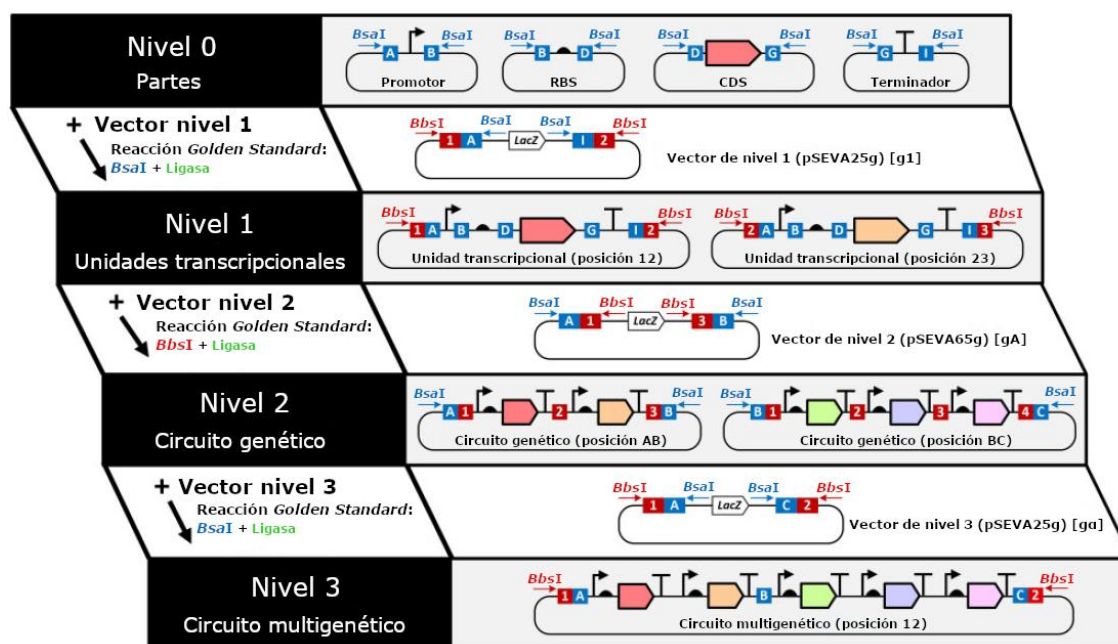


Figura R1.1. Esquema del ensamblaje secuencial del *Golden Standard*. Las partes o niveles 0 se ensamblan en un vector de nivel 1 para formar una unidad transcripcional. A su vez, varias unidades transcripcionales se pueden ensamblar en un vector de nivel 2 y formar un circuito genético. Los circuitos multigenéticos se construyen en un vector de nivel 3. Todas las reacciones de ensamblaje consisten en varios ciclos de digestión por enzimas de restricción de tipo IIS y ligación. Adaptada de Blázquez et al. (2023).

1.3.1. Construcción de la colección de plásmidos con *ori* RSF1010

Para conseguir ampliar el GS, se decidió trabajar con el origen de replicación RSF1010 que se caracteriza por ser de amplio rango de hospedador debido a que no depende de la maquinaria de replicación del hospedador ya que contiene los genes *repA*, *repB* y *repC* (Frey et al., 1992). Esto permite que los plásmidos sean replicativos tanto en bacterias Gram negativas como Gram positivas y, entre ellas, en micobacterias, como se ha demostrado previamente (García-Fernández et al., 2014).

La colección de plásmidos pSEVA ya incluye plásmidos con este origen de replicación RSF1010 (e.g., pSEVA25X, pSEVA65X), pero requieren ser modificados para generar una colección de plásmidos que puedan ser utilizados en GS. Primero, los plásmidos deben ser ‘domesticados’, es decir, hay que eliminar todas las dianas de *BbsI* y *BsaI* presentes en el vector. En concreto, la secuencia de este *ori* presenta dos dianas de *BbsI*: una en el gen *repA* y otra en la región intergénica entre *repA* y *repB* (Fig. R1.2). Ambas dianas se eliminaron mediante mutagénesis dirigida, amplificando por PCR con los pares de oligonucleótidos IPM235-IPM194 e IPM196-IPM197 (Tabla M3) dos fragmentos divergentes y ligando por *Golden Gate*. De esta forma se obtuvieron los plásmidos curados pSEVA25g (KmR) y pSEVA65g (GmR).

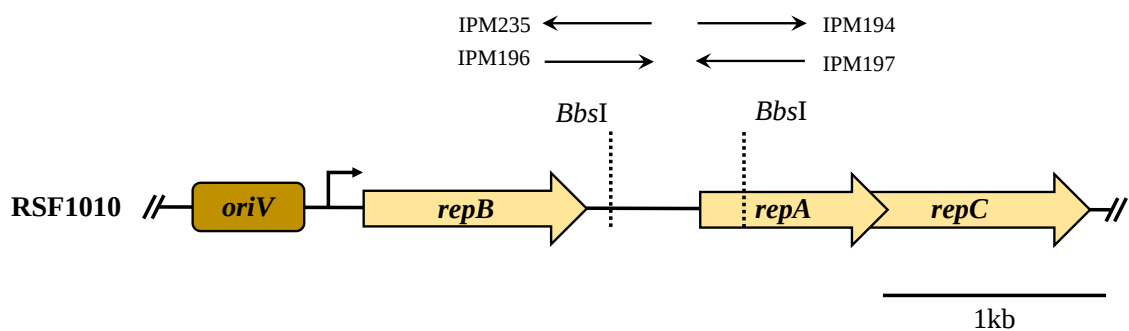


Figura R1.2. ‘Domesticación’ del *ori* RSF1010 mediante PCR. Las dos dianas presentes en el *ori* RSF1010 se eliminaron introduciendo dos mutaciones silenciosas en las dianas de restricción.

A continuación, se insertaron en el sitio de clonaje múltiple de los vectores las dianas de *BbsI* y *BsaI* con sus FS necesarios para permitir el clonaje modular y jerárquico de las unidades transcripcionales (Fig. R1.3). Entre las dianas de *BbsI* y *BsaI*, además, se insertó el gen *lacZα* de forma que permite la selección de las células transformantes por colonias blancas y azules en presencia de X-Gal e IPTG. De esta forma, se obtuvieron 16 vectores que constituyen la colección de plásmidos con *ori* RSF1010 para su uso en GS. En la colección encontramos 7 vectores de nivel 1 que se pueden combinar y clonar a su vez en los 5 vectores de nivel 2 diseñados. Potencialmente, estos niveles 2 se pueden combinar en los 4 vectores de nivel 3 contruidos (Fig. R1.3). La correcta construcción de los nuevos plásmidos se comprobó por secuenciación mediante nanoporo en Plasmidsaurus Inc.

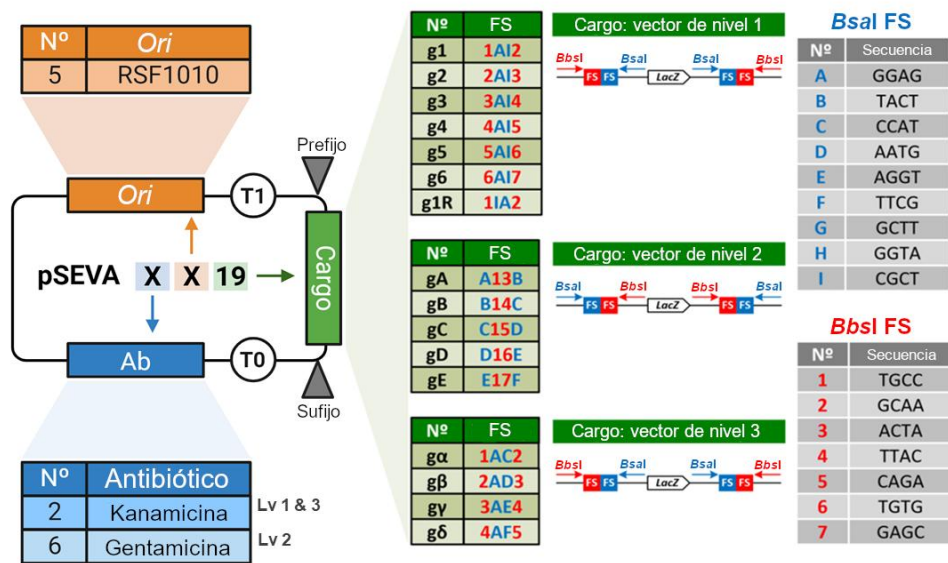


Figura R1.3. Colección del Golden Standard de plásmidos con ori RSF1010. Esta colección presenta 7 vectores de nivel 1 (pSEVA25g1-g6 y pSEVA25g1Reverse), 5 vectores de nivel 2 (pSEVA65gA-E) y 4 vectores de nivel 3 (pSEVA25gα-δ). Todos ellos presentan en el cargo el gen *lacZ* flanqueado por las dianas de *BbsI* y *BsaI*, cada uno con unos sitios de fusión específicos para permitir el clonaje modular por complementariedad. Adaptada de Blázquez et al. (2023).

1.3.2. Validación de la nueva colección de *Golden Standard* en *M. smegmatis*

1.3.2.1. Caracterización de la colección de promotores

Aprovechando la modularidad del GS y para ensayar la actividad de distintos promotores constitutivos en *M. smegmatis*, se seleccionaron aquellos utilizados por Blázquez et al. (2023) y otros específicos de micobacterias: los promotores J23100, J23102 y J23106 de la librería de Anderson (Iverson et al., 2016); Pem7, BG17 y BG37 de la librería de Zobel (Tiso et al., 2016) y Phsp60 (Stover et al., 1991), Pimyc (Chang et al., 2009) y PhspX (Oliveira et al., 2019) como promotores específicos del género. Estos promotores se clonaron en vectores de nivel 0 y se ensamblaron en el plásmido de nivel 1 pSEVA25g para formar una unidad transcripcional. Para ello, se fijaron el RBS estándar (AGGAGG), la *rfp*, que codifica la proteína roja fluorescente, como gen reportero y el terminador sintético B1006, como se esquematiza en la Figura R1.4. De esta forma se obtuvieron los plásmidos p25_J23100_RFP, p25_J23102_RFP, p25_J23106_RFP, p25_Pem7_RFP, p25_BG17_RFP, p25_BG37_RFP, p25_Phsp60_RFP, p25_PhspX_RFP y p25_Pimyc_RFP (Tabla M2). Estos plásmidos se transformaron en *M. smegmatis* y *E. coli* y se ensayaron en un lector de placas Victor Nivo PerkinElmer en placas multipocillo, donde se monitorizó el crecimiento (DO₆₀₀) y la fluorescencia como se describe en el apartado 9 de Materiales y métodos. Como control se construyó el plásmido p25_linker_RFP (Tabla M2) en el que el promotor se sustituyó por una secuencia de ADN denominada *linker* que no presenta actividad promotora. Los

resultados obtenidos se analizaron con el *software* QurvE, que permite analizar crecimiento y fluorescencia respecto al tiempo.

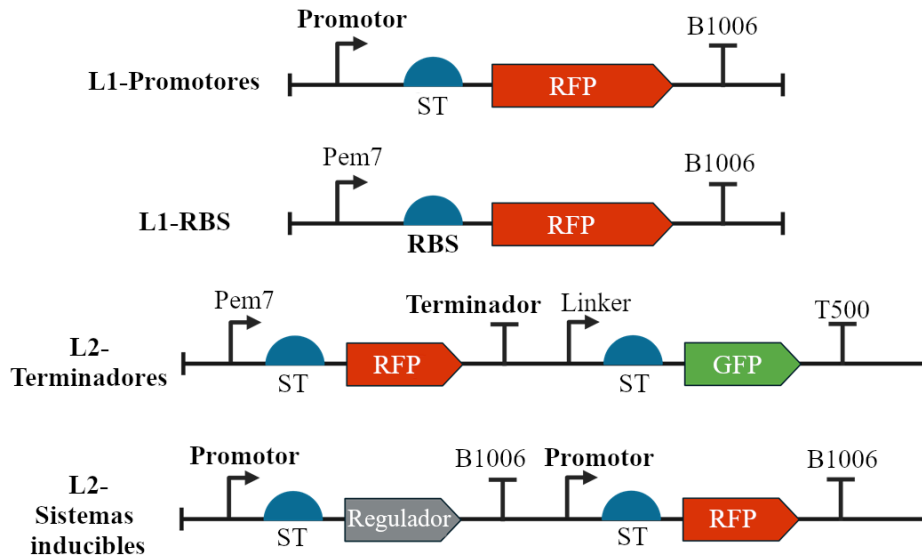


Figura R1.4. Esquema de las construcciones realizadas para ensayar la actividad de los promotores, RBS, terminadores y sistemas inducibles. RBS: *Ribosome binding site*; ST: RBS estándar; RFP: proteína roja fluorescente; GFP: proteína verde fluorescente.

La Figura R1.5A muestra que, en el caso de *E. coli*, la mayor intensidad de fluorescencia observada se debe a los promotores de la colección de Zobel (Pem7, BG17 y BG37), alcanzando los $3 \cdot 10^7$ mRFP/DO h⁻¹ de unidades relativas de fluorescencia (URF). Por el contrario, la señal emitida de los promotores de la colección de Anderson fue menor (entre $1\text{-}5 \cdot 10^7$ mRFP/DO h⁻¹). Estos datos correlacionan con los observados por Blázquez et al (2023), en los que los promotores obtenidos de la colección de Zobel, ensayados en plásmidos con el *ori* RK2, presentaron una mayor actividad que aquellos de la colección de Anderson en las cepas ensayadas. Por otro lado, la mayor intensidad de fluorescencia en *M. smegmatis* se observó con las construcciones que portaban los promotores J23100, J23102, BG17 y BG37 (Fig. R1.5B). Sorprendentemente, estos valores distan mucho de los observados por los promotores Pem7, Phsp60, PhspX y Pimyc, de los cuales 3 de ellos son específicos de micobacterias y cuya señal no superó los $3 \cdot 10^7$ mRFP/DO h⁻¹. Gracias a este trabajo se dispone de 8 promotores que presentan un amplio rango de actividades en *M. smegmatis*, lo que ofrece una gran versatilidad en el uso de GS en esta bacteria.

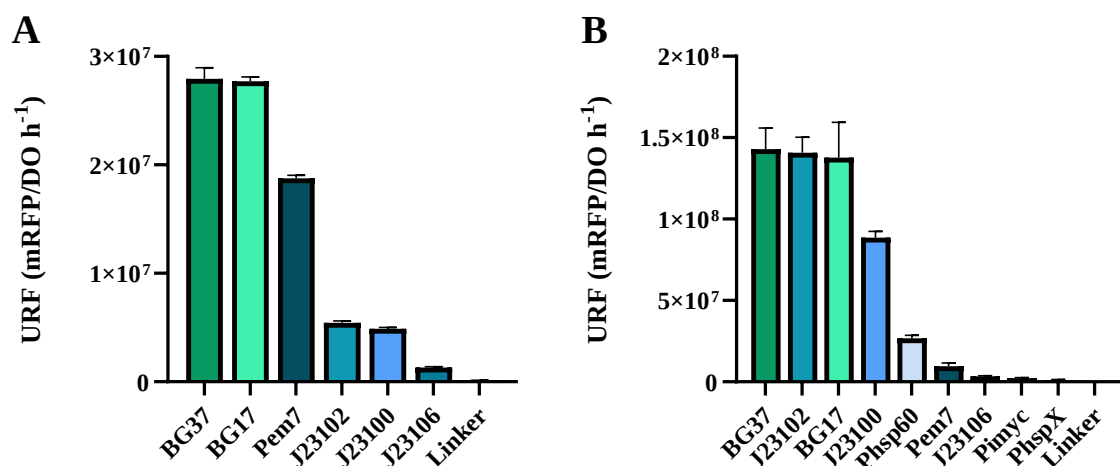


Figura R1.5. Análisis de la actividad de distintos promotores en *E. coli* (A) y *M. smegmatis* (B). La potencia de cada promotor se representa como la señal de fluorescencia emitida normalizada por la DO₆₀₀ por hora. URF: unidades relativas de fluorescencia. Los valores corresponden a la media de 3 réplicas técnicas.

1.3.2.2. Caracterización de la colección de RBSs

En la naturaleza, los ARNm pueden ser monocistrónicos con un RBS y una secuencia codificante, o pueden ser bi- o policistrónicos con dos o más secuencias codificantes, cada una con su propio RBS (Abdi Ghavidel et al., 2022). En función de la secuencia RBS y de la estructura mono-, bi- o policistrónica del ARNm, se pueden obtener niveles variables de traducción. Para comprobar la actividad de algunos de los RBS presentes en la colección GS en *M. smegmatis*, se construyeron varios plásmidos de nivel 1 fijando el promotor Pem7, el gen reportero *rfp* y el terminador B1006 (Fig. R1.4). De entre los RBS monocistrónicos se probaron en estándar (ST) y el B0034 y de los bicistrónicos, en los que un pequeño polipéptido de 17 aminoácidos forma el primer cistron, el BCD12 y BCD2 (Blázquez et al., 2023). De esta forma se obtuvieron los plásmidos p25_ST_RFP, p25_B0034_RFP, p25_BCD12_RFP y p25_BCD2_RFP (Tabla M2). Como control se ensambló el plásmido p25_Pem7_Linkers_RFP (Tabla M2) en el que el RBS se sustituyó por un *linker*. Estas construcciones se transformaron tanto en *E. coli* como en *M. smegmatis* y se ensayaron como se ha indicado anteriormente.

Los RBS bicistrónicos presentaron unos niveles de traducción mayores en *E. coli* y, por tanto, se obtuvieron niveles más altos de fluorescencia, siendo 3 veces mayor que lo observado con el RBS ST (Fig. R1.6A). Una vez más estos datos son similares a los observados por Blázquez et al. (2023) ensayados con el *ori* RK2. En el caso de *M. smegmatis*, la mayor expresión se observó con el RBS BCD2, siendo casi 2 veces mayor que la observada con B0034 y BCD12 y 3 veces mayor que la observada con el RBS ST (Fig. R1.6B). Las construcciones bicistrónicas, por tanto, son las que mejor resultado han dado en ambas cepas, llegando a niveles máximos de fluorescencia bastante similares en *M. smegmatis* y *E. coli* (3,6 y 6 · 10⁷ mRFP/DO h⁻¹, respectivamente). Esto demuestra la

gran aplicabilidad que tienen los ARNm bicistrónicos tanto en bacterias Gram negativas como en Gram positivas.

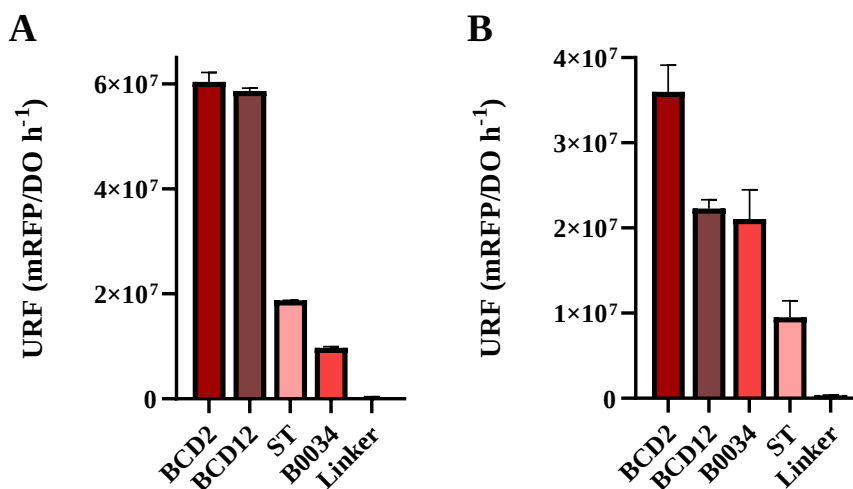


Figura R1.6. Caracterización de la actividad de distintos RBS en *E. coli* (A) y *M. smegmatis* (B). La fuerza de cada RBS se representa como la señal de fluorescencia emitida normalizada por la DO₆₀₀ por hora. URF: unidades relativas de fluorescencia. Los valores corresponden a la media de 3 réplicas técnicas.

1.3.2.3. Caracterización de la colección de terminadores de la transcripción

Los terminadores de la transcripción presentes en la colección incluyen los terminadores naturales *rnpB* y *rpoC* y los terminadores sintéticos T500, B1006 y B0014 (Blázquez et al., 2023). Para ensayar su efecto en *M. smegmatis*, se construyeron 5 plásmidos de nivel 2 que se componían de las siguientes partes: el promotor Pem7, el RBS ST, el gen *rfp*, el terminador a caracterizar, otro RBS ST adicional, el gen *gfp*, que codifica la proteína verde fluorescente, y el terminador T500 (Fig. R1.4). De esta forma, el gen reportero *rfp* se expresa bajo el promotor constitutivo Pem7 y, en función de la eficacia del terminador, se detectará más o menos expresión del siguiente gen reportero *gfp*. Al calcular la ratio de ambas señales se puede conocer así el efecto del terminador.

Se obtuvieron así los plásmidos p65_RFP_B0014_GFP, p65_RFP_B1006_GFP, p65_RFP_rpoC_GFP, p65_RFP_rnpB_GFP, p65_RFP_T500_GFP y p65_RFP_Linkers_GFP (Tabla M2). El plásmido control presenta un *linker* en el sitio del terminador, de forma que ambos genes reporteros se expresan por igual. Todas las construcciones se transformaron en *E. coli* y *M. smegmatis* y se ensayaron como está descrito en el apartado 9 de Materiales y métodos. En el caso de *E. coli*, salvo el terminador sintético T500 cuya eficiencia fue del 38%, el resto presentó una actividad entre 74-96%, siendo el mejor de ellos el terminador natural *rpoC*, seguido del terminador sintético B1006 (Fig. R1.7A). Estos valores no se observaron sin embargo en *M. smegmatis*, donde la mejor eficiencia oscila entre 50-75%, siendo el mejor de ellos el

terminador sintético B1006 (Fig. R1.7B). En esta cepa además el terminador B0014 no tuvo ningún efecto, mientras que el terminador *rpoC* mostró una eficiencia del 24%, en contraste con lo que sucede en *E. coli*.

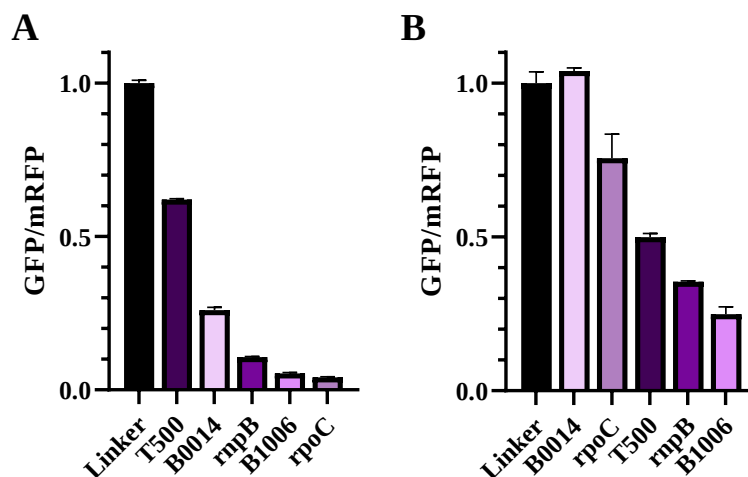


Figura R1.7. Estudio de la capacidad terminadora de distintos terminadores en *E. coli* (A) y *M. smegmatis* (B). La fuerza de cada terminador se calculó haciendo la ratio entre las señales de fluorescencia emitidas por la GFP y la mRFP. Los valores corresponden a la media de 3 réplicas técnicas.

1.3.2.4. Caracterización de los sistemas inducibles

Los sistemas inducibles son una herramienta muy utilizada en la biología molecular, ya que principalmente permiten controlar el tiempo y los niveles de expresión de un gen, aportando así una mayor versatilidad en distintos procesos celulares. Estos sistemas se basan principalmente en la expresión constitutiva de un regulador transcripcional como un represor (LacI, RhaSR, XylS) y la expresión de un gen de interés bajo un promotor (Ptrc, PrhaBAD, Pm) que responda a este regulador transcripcional. Los 3 sistemas de inducción presentes en el GS son los pares LacI/Ptrc, RhaSR/PrhaBAD y XylS/Pm, inducibles por IPTG, ramnosa y ácido *m*-toluico, respectivamente. Para su aplicación en GS, primero se construyeron 2 plásmidos de nivel 1: uno con un promotor constitutivo y el represor y otro con el promotor inducible y el gen reportero *rfp*. Con ambos plásmidos se construyó un plásmido de nivel 2 en los que la expresión del gen *rfp* se encuentra controlada por el promotor inducible (Fig. R1.4), obteniéndose así las construcciones p65_LacI_Ptrc_RFP, p65_RhaSR_PrhaBAD_RFP y p65_XylS_Pm_RFP (Tabla M2). Una vez transformadas en *E. coli* y *M. smegmatis*, se ensayó el efecto del inductor con distintas concentraciones (desde 100 mM hasta 1 μ M) probando diluciones 10^{-1} .

Tal y como se observa en la Figura R1.8A, todas las construcciones presentaron represión e inducción en *E. coli*. La fluorescencia basal en todos los sistemas en ausencia del inductor no superó los $1,4 \cdot 10^6$ mRFP/DO h $^{-1}$. La concentración de IPTG a la que se

observó la mayor inducción fue de 100 μM , mientras que para la ramnosa y el ácido *m*-toluico fue de 10 mM, generando en los 3 casos unos niveles de fluorescencia un orden de magnitud por encima de los niveles basales ($2,4\text{-}5,6 \cdot 10^7$ mRFP/DO h^{-1}). Sin embargo, ninguno de los 3 sistemas de inducción funcionó en *M. smegmatis*, ya que en todos ellos no se observó un aumento de los niveles de fluorescencia en presencia del inductor independientemente de la concentración de éste (Fig. R1.8B). En el caso del sistema LacI/Ptrc, las células presentaban una coloración rosada en ausencia de IPTG, sugiriendo una elevada expresión basal y/o una ausencia de represión. Sin embargo, en los sistemas RhaSR/PrhaBAD y XylS/Pm la expresión basal fue muy baja y tampoco se observó un aumento de fluorescencia en presencia del inductor. Esto podría deberse a que el promotor inducible no está siendo reconocido por la ARN polimerasa de *M. smegmatis*.

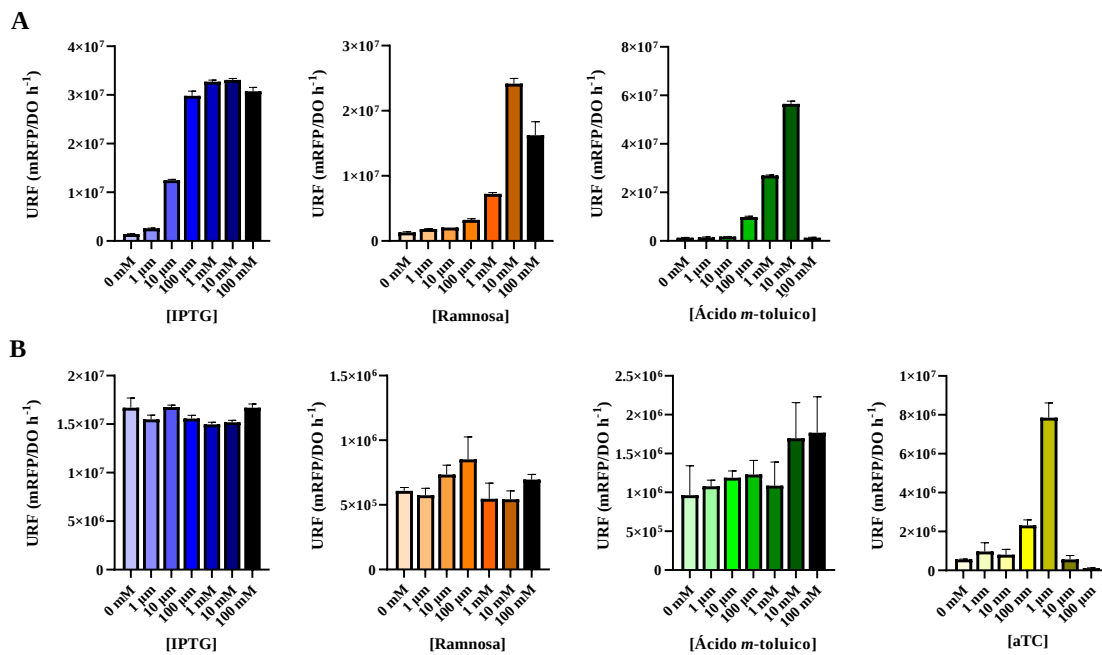


Figura R1.8. Caracterización de distintos sistemas inducibles en *E. coli* (A) y en *M. smegmatis* (B). Las construcciones ensayadas corresponden a los sistemas LacI/Ptrc (azul), RhaSR/PrhaBAD (naranja), XylS/Pm (verde) y TetR/Ptet (amarillo). El nivel de inducción se estudió según la fluorescencia emitida normalizada por la DO₆₀₀ por hora en función de la concentración de inductor añadido. URF: unidades relativas de fluorescencia. Los valores corresponden a la media de 3 réplicas técnicas.

Ante la falta de un sistema de inducción funcional en *M. smegmatis*, se decidió construir uno desde cero para ampliar así la colección de elementos génicos del GS. Así se decidió replicar el sistema TetR/Ptet, empleado previamente por Arnold et al. (2018) en *M. smegmatis*, basado en el represor TetR y cuyo inductor es la anhidrotetraciclina (aTC). El gen *tetR* y el promotor Ptet se clonaron independientemente en plásmidos de nivel 0 para poder ser utilizados en las posteriores construcciones. El sistema inducible se construyó en un plásmido de nivel 2 donde el gen *tetR* se clonó junto al promotor débil Pimyc, el RBS ST y el terminador B1006, formando una unidad transcripcional, y el gen

reportero *rfp* se clonó junto al promotor inducible Ptet, el RBS ST y el terminador B1006, obteniendo así el plásmido p65_TetR_Ptet_RFP (Tabla M2). Este plásmido se transformó en *M. smegmatis* y se ensayó probando un rango de concentraciones desde 0,1 mM hasta 1 nM. Como se puede observar en la Figura R1.8, con este sistema sí se observó tanto represión en ausencia de aTC como inducción en presencia de aTC, siendo 1 μ M la concentración en la que se obtuvo la mayor señal de fluorescencia. Sin embargo, la desventaja que presenta este sistema de inducción en *M. smegmatis* es su bajo nivel de expresión, ya que es 3 veces menor que la potencia del promotor constitutivo débil Phsp60 (Fig. R1.5), lo que limita su uso.

Cabe destacar que, en algunos casos, las concentraciones de los inductores utilizadas en el ensayo resultaron ser tóxicas para los microorganismos. No se observó crecimiento cuando *E. coli* se cultivó con 10 o 100 mM de ramnosa o 100 mM de ácido *m*-toluico, mientras que en *M. smegmatis*, no hubo crecimiento celular en presencia de 100 μ M de aTC o de 10 o 100 mM de ácido *m*-toluico. Sin embargo, en esta cepa no se detectó toxicidad a ninguna concentración de ramnosa o de IPTG.

1.4. Producción bacteriana de progesterona

La progesterona es una hormona sexual con aplicaciones farmacéuticas que ha sido nominada como un fármaco crucial por la OMS, por lo que constituye una molécula relevante para la industria esteroidea. Como se ha descrito previamente, la progesterona se puede producir desde la sapogenina diosgenina mediante 8 pasos de síntesis química o desde la sintona esteroidea C22 4-HBC y solo 2 pasos de síntesis química (Peng et al., 2021). Por tanto, estos procesos siguen dependiendo del uso de distintos reactivos químicos que pueden ser contaminantes en el medio ambiente.

En los seres humanos la progesterona se sintetiza a partir de colesterol mediante la actuación del citocromo mitocondrial CYP11A1 (también denominado CYP450scc), sin embargo, hasta hace poco no se conocían citocromos similares en bacterias. Hoy se sabe que existen 2 CYP450scc descritos en la patente EP2386634A1 (Yamagishi et al., 2011) que provienen de las bacterias *N. aromaticivorans* y *N. subterraneum* y que presentan actividad colesterol hidroxilasa/C20-22 liasa. Estas enzimas catalizan dos hidroxilaciones consecutivas en las posiciones C20 y C22 de la cadena lateral del colesterol y de la colestenoína y acaban cortando su cadena alifática, lo que genera pregnenolona y progesterona, respectivamente.

Para producir progesterona directamente desde esteroides mediante un proceso de biotransformación sin la necesidad de pasos de síntesis química, se decidió aplicar las nuevas herramientas de GS desarrolladas en el apartado anterior para expresar uno de estos CYP450scc bacterianos en *M. smegmatis*. Como se ha mencionado previamente, esta cepa es utilizada a nivel industrial como factoría celular para la síntesis de sintonas esteroideas y presenta además una serie de ventajas, ya que tanto el transporte como el metabolismo de esteroides se encuentra muy estudiado (Galán et al., 2017; García-Fernández et al., 2017b). En esta estrategia, el colesterol (o en su caso los fitoesteroides)

sería utilizado como sustrato por los CYP450scc bacterianos, que romperían su cadena lateral formando pregnenolona. Esta pregnenolona sería después oxidada a progesterona por la 3 β -HSD involucrada en el metabolismo de los esteroides en *M. smegmatis* (Fig. R1.9). Para evitar que la cadena del colesterol se degrade por el mecanismo propio, se requiere la eliminación previa de los CYP450 encargados de comenzar la degradación de la cadena lateral del colesterol, como se ha descrito en el apartado 2.1.1 de Introducción.

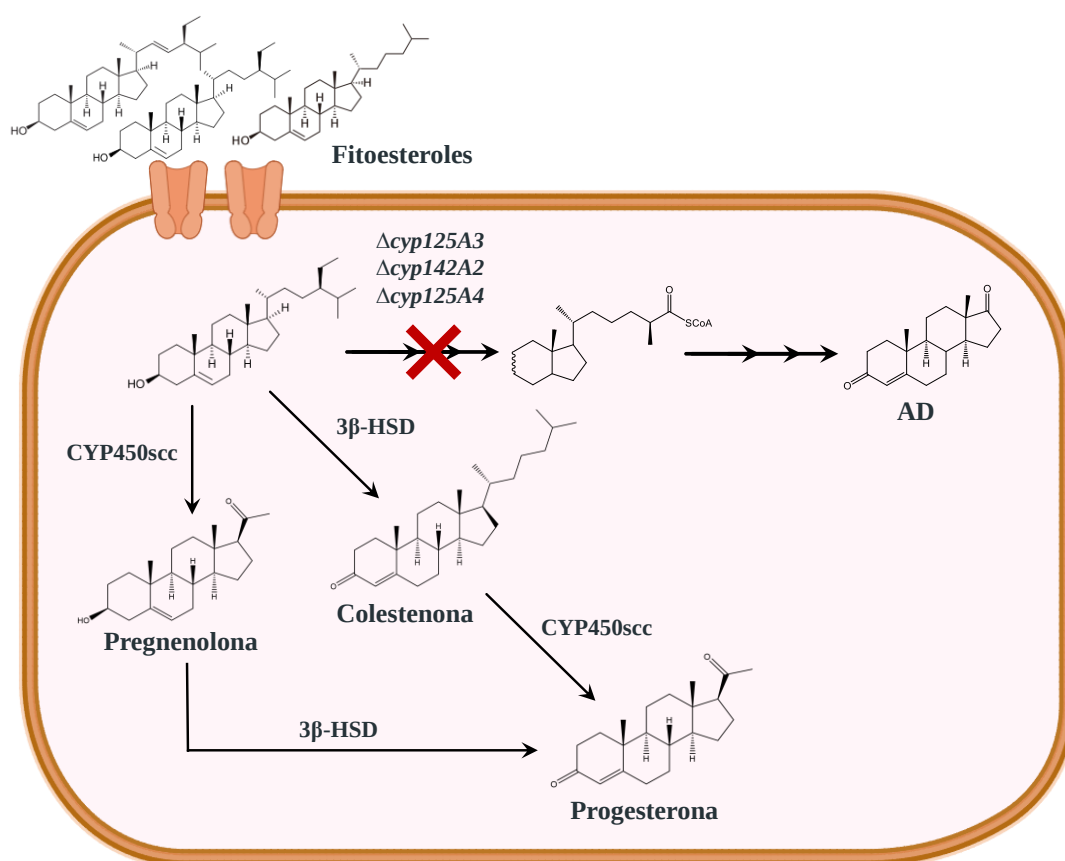


Figura R1.9. Esquema de la ruta sintética propuesta para la producción de progesterona en *M. smegmatis*. 3 β -HSD: 3 β -hidroxiesteroide deshidrogenasa; AD: androstenediona.

1.4.1. Construcción de una cepa de *M. smegmatis* deficiente en el consumo de esteroides

La cepa silvestre de *M. smegmatis* no es capaz de utilizar la progesterona como fuente de carbono y de energía, pero sí presenta dos regulones de 83 y 15 genes catabólicos y varios genes aislados especializados en la degradación de los esteroides (Uhía et al., 2011). Es por ello que, para poder utilizar los esteroides como materia de partida para producir la hormona de interés, en primer lugar, hay que generar un mutante que no modifique los esteroides. En este sentido, García-Fernández et al. (2013) identificaron dos CYP450 que inician la degradación del colesterol activando la cadena lateral y obtuvieron un doble mutante de *M. smegmatis* llamado $\Delta Cyp125\Delta Cyp142$, que presentó una mayor

fase *lag* respecto a la cepa silvestre cuando se cultivaba en presencia de colesterol, sugiriendo una posible redundancia enzimática en estos CYP450 degradadores de colesterol. Frank et al. (2015) describieron que el CYP125A4 de *M. smegmatis*, codificado en el gen *MSMEG_3524*, es capaz de unir y actuar (aunque débilmente) sobre sustratos esteroideos tales como el colesterol y la colesteno, por lo que fue propuesto como el tercer CYP450 que compensa la ausencia de los dos citocromos CYP125A3 y CYP142A2 de la ruta de degradación y permite que el doble mutante termine creciendo en colesterol.

Teniendo en cuenta estos datos, la cepa $\Delta Cyp125\Delta Cyp142$ se utilizó como punto de partida para deletar sobre ella el gen *MSMEG_3524* del CYP125A4 para conseguir un mutante que no fuera capaz de crecer en colesterol como única fuente de carbono y energía. De esta forma, la cepa se transformó con el plásmido suicida pJQ3524 (Tabla M2), que contiene un fragmento UP y otro DOWN homólogos a las regiones aguas arriba y aguas abajo de *MSMEG_3524*, y se seleccionaron primero las células que recombinaron e insertaron el plásmido y después aquellas que realizaron el segundo paso de recombinación, como está descrito en el apartado 4.4 de Materiales y métodos. Así se consiguió el mutante triple $\Delta Cyp125A3-\Delta 142A2-\Delta 125A4$. Esta nueva cepa se cultivó a continuación en medio mínimo líquido con colesterol 1,8 mM como única fuente de carbono y de energía. Tal y como muestra la Figura R1.10, el triple mutante continúa conservando capacidad para crecer en presencia de colesterol. Además, no se observa una mejora significativa respecto a la cepa mutante de partida $\Delta Cyp125\Delta Cyp142$. Este resultado sugiere que, a diferencia de lo propuesto por Frank et al. (2015), el CYP125A4 no es responsable de activar la cadena lateral del colesterol para su degradación *in vivo* y por tanto existe un CYP450 adicional en el genoma de *M. smegmatis* capaz de catalizar esta reacción. A pesar de ello, se decidió continuar con la cepa triple mutante $\Delta Cyp125A3-\Delta 142A2-\Delta 125A4$ como factoría celular para producir progesterona aprovechando la amplia ventana de casi 48 h que proporciona la fase *lag* observada cuando crece con colesterol.

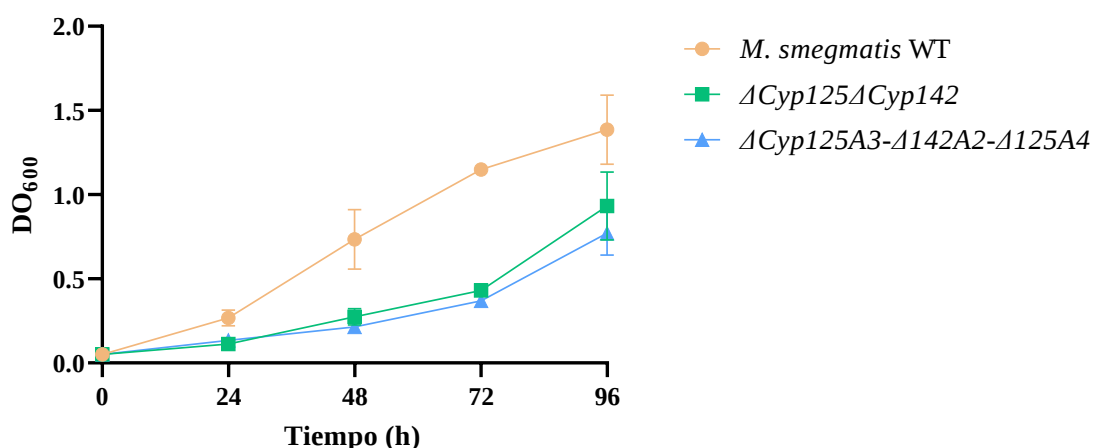


Figura R1.10. Curva de crecimiento de las cepas *M. smegmatis* WT, $\Delta Cyp125\Delta Cyp142$ y $\Delta Cyp125\Delta 3-1142\Delta 2-1125\Delta 4$ en presencia de colesterol 1,8 mM. Los valores representan la media de 3 réplicas biológicas.

1.4.2. Biotransformación de esteroides en progesterona en *M. smegmatis*

Una vez obtenida la cepa mutante $\Delta Cyp125\Delta 3-1142\Delta 2-1125\Delta 4$ que se utilizará como chasis celular para realizar las biotransformaciones, se procedió a construir los plásmidos necesarios para expresar el CYP450_{scc} de *N. aromaticivorans* con las proteínas acompañantes ferredoxina (Fdx) y ferredoxina reductasa (FdxR), necesarias para producir progesterona a partir de esteroides.

Estos CYP450_{scc} bacterianos se clasifican como CYP450 de clase I, y se caracterizan por depender de dos proteínas auxiliares: la Fdx o adrenoxina (Adx) y sus respectivas reductasas FdxR o AdxR, que les proporcionan los electrones necesarios para catalizar la monooxigenación (Schultes et al., 2024; Strushkevich et al., 2011; Yamagishi et al., 2011). En ocasiones se desconoce el par Fdx/FdxR específico de cada CYP450 ya que el mismo par puede actuar con distintos citocromos, llegando a poder ser reemplazada su actividad incluso por proteínas de otros microorganismos, como es el caso de la putidaredoxina y la putidaredoxina reductasa (par CamAB) (Koga et al., 1989). En el caso del CYP450_{scc} descrito en la patente EP2386634A1 (Yamagishi et al., 2011) y perteneciente a *N. aromaticivorans*, esta enzima se encuentra codificada por el gen *Saro_3659* y a continuación presenta una Fdx codificada por el gen *Saro_3658*, lo que sugiere que ambas proteínas actúan en conjunto. Por otro lado, Yang et al. (2010) caracterizaron la ferredoxina reductasa ArR de *N. aromaticivorans* como sistema de transferencia de electrones, confirmando su actividad con varios CYP450s de clase I. Por esta razón los genes *Saro_3659* (CYP450_{scc}), *Saro_3658* (Fdx) y *Saro_0216* (ArR) fueron los elegidos para producir progesterona en *M. smegmatis*.

Para comprobar el papel del CYP450_{scc}, de la Fdx y de la ArR en *M. smegmatis*, se construyeron varios plásmidos conteniendo solo el CYP450_{scc}, el CYP450_{scc} con la Fdx o el CYP450_{scc} y la Fdx y la ArR. Para ello se aprovecharon las nuevas herramientas genéticas de GS construidas *ad hoc* para micobacterias. Como una primera aproximación, se decidió que cada gen se expresara en una unidad transcripcional y se fijó para todos los genes el promotor Phsp60, el RBS ST y el terminador B1006. De esta forma se construyeron los plásmidos p25_Phsp60_CYP, p65_Phsp60_CYP_Fdx y p65_Phsp60_CYP_Fdx_ArR (Tabla M2), que se transformaron en la cepa $\Delta Cyp125\Delta 3-1142\Delta 2-1125\Delta 4$ obteniendo así los transformantes $\Delta Cyp125\Delta 3-1142\Delta 2-1125\Delta 4$ (p25_Phsp60_CYP), $\Delta Cyp125\Delta 3-1142\Delta 2-1125\Delta 4$ (p65_Phsp60_CYP_Fdx) y $\Delta Cyp125\Delta 3-1142\Delta 2-1125\Delta 4$ (p65_Phsp60_CYP_Fdx_ArR). Para obtener la cepa control se transformó con el plásmido vacío pSEVA65g (Tabla M2) obteniendo la cepa $\Delta Cyp125\Delta 3-1142\Delta 2-1125\Delta 4$ (pSEVA65g).

Aunque el triple mutante $\Delta Cyp125A3-\Delta 142A2-\Delta 125A4$ presenta fase *lag* muy extensa, la producción de progesterona a partir de colesterol se ensayó con células en reposo o *resting cells* para reducir el tiempo de la biotransformación, ya que así se favorece que el colesterol sea utilizado como materia prima y que no entre en el catabolismo de la bacteria. Las cepas se cultivaron primero en medio rico con el antibiótico correspondiente y en presencia de 1 mM del ácido 5-aminolevulínico y de 1 mM de $FeSO_4$ para favorecer la producción del grupo hemo, necesario para la actividad del CYP450_{scc}. Una vez crecidas, las células se concentraron hasta una $DO_{600}=18$ y se añadieron al medio de biotransformación, que contiene colesterol 500 mg/L y glicerol 2% (p/v) que se añadió como fuente de energía. El medio se analizó a las 30 h de reacción para detectar la presencia de progesterona. La Figura R1.11 muestra el cromatograma de HPLC de la cepa $\Delta Cyp125A3-\Delta 142A2-\Delta 125A4$ (p65_Phsp60_CYP_Fdx_ArR) a las 0 h y a las 30 h de la biotransformación. Se observa la acumulación en el tiempo de progesterona como producto y la acumulación de la versión deshidrogenada de la progesterona 1,2-dehidroprogesterona, lo que se debe muy posiblemente a la acción de las KstD nativas de *M. smegmatis* sobre la progesterona.

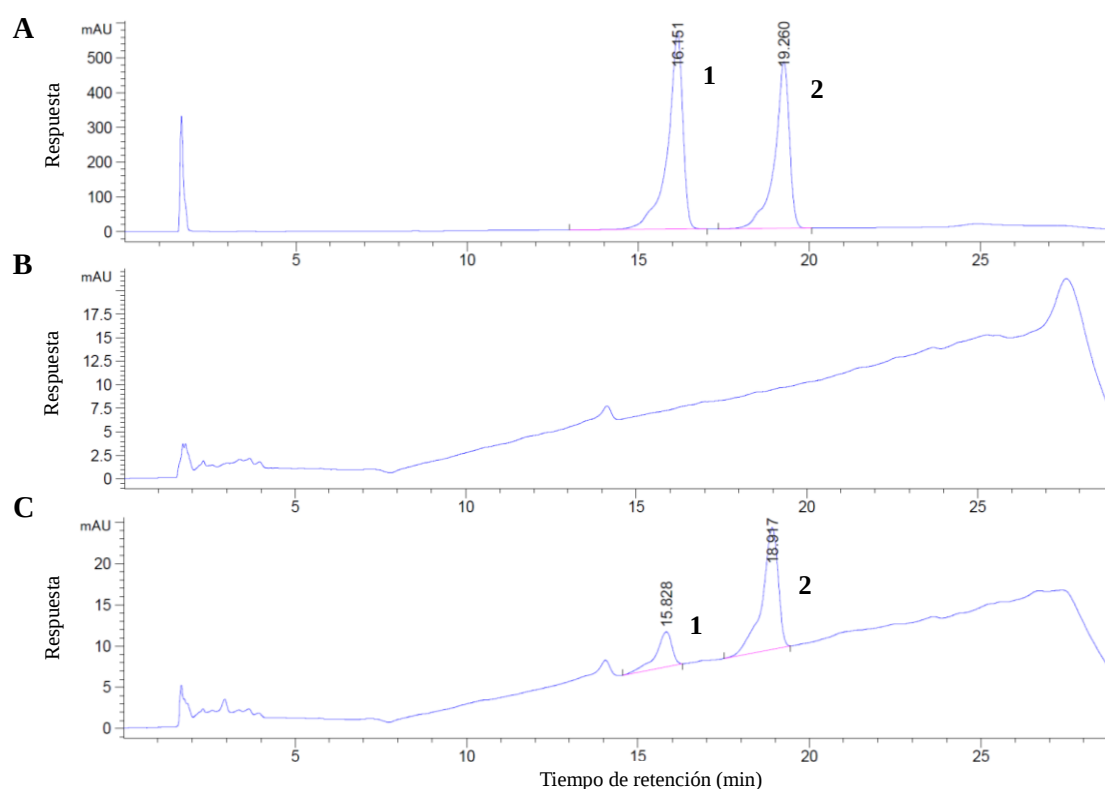


Figura R1.11. Análisis por HPLC de los esteroides producidos por la cepa $\Delta Cyp125A3-\Delta 142A2-\Delta 125A4$ (p65_Phsp60_CYP_Fdx_ArR). Se compara el cromatograma (A) de los patrones de 1,2-dehidroprogesterona (pico 1) y de progesterona (pico 2) con los cromatogramas obtenidos de la biotransformación de 500 mg/L de colesterol a las 0 h (B) y a las 30 h (C).

Se cuantificó la producción de progesterona de las distintas cepas observándose que la cepa $\Delta Cyp125A3-\Delta142A2-\Delta125A4$ (p25_Phsp60_CYP), que expresa solo el CYP450scc, acumuló 0,64 mg/L de progesterona (Fig. R1.12A). Esta producción aumentó hasta los 1,38 mg/L cuando se expresaron a la vez el CYP450scc y la Fdx en la cepa $\Delta Cyp125A3-\Delta142A2-\Delta125A4$ (p65_Phsp60_CYP_Fdx), pero el mejor valor se observó cuando se utilizó la cepa $\Delta Cyp125A3-\Delta142A2-\Delta125A4$ (p65_Phsp60_CYP_Fdx_ArR), que acumuló hasta 2,67 mg/L de la hormona de interés. Solo las cepas $\Delta Cyp125A3-\Delta142A2-\Delta125A4$ (p65_Phsp60_CYP_Fdx) y $\Delta Cyp125A3-\Delta142A2-\Delta125A4$ (p65_Phsp60_CYP_Fdx_ArR) acumularon también 1,2-dehidroprogesterona a las 30 h, llegando a una concentración de 0,27 y 0,54 mg/L, respectivamente (Fig. R1.12B).

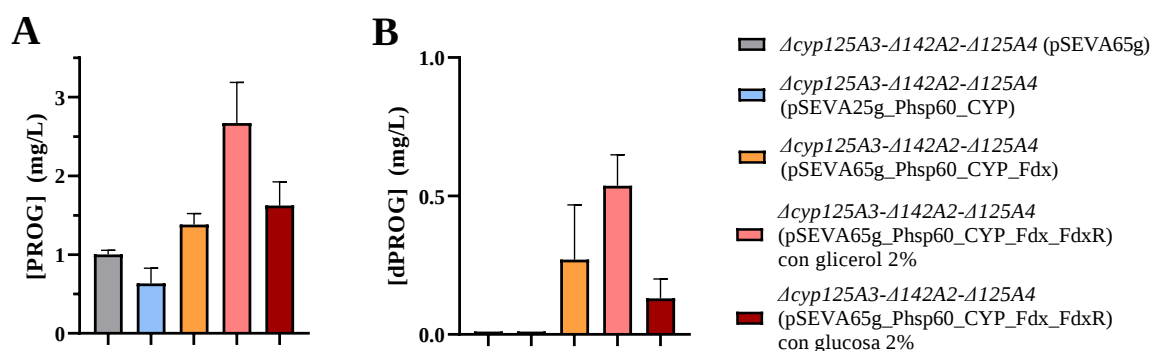


Figura R1.12. Producción de progesterona (A) y de 1,2-dehidroprogesterona (B) a partir de colesterol por las cepas $\Delta Cyp125A3-\Delta142A2-\Delta125A4$ (pSEVA65g), $\Delta Cyp125A3-\Delta142A2-\Delta125A4$ (p25_Phsp60_CYP), $\Delta Cyp125A3-\Delta142A2-\Delta125A4$ (p65_Phsp60_CYP_Fdx) y $\Delta Cyp125A3-\Delta142A2-\Delta125A4$ (p65_Phsp60_CYP_Fdx_ArR) tras 30 h de biotransformación. La cepa $\Delta Cyp125A3-\Delta142A2-\Delta125A4$ (p65_Phsp60_CYP_Fdx_ArR) se ensayó con glicerol o glucosa 2% (p/v). Los valores representan la media de 3 réplicas biológicas.

Estos datos confirman que el sistema CYP450scc-Fdx-ArR es funcional en *M. smegmatis* y que, como el resto de CYP450 de clase I, el CYP450scc requiere de la acción de la Fdx y de la ArR. Por tanto, se seleccionó la cepa que portaba el plásmido p65_Phsp60_CYP_Fdx_ArR para probar distintas condiciones de biotransformación y tratar de optimizar la producción de progesterona. Cabe destacar que el control $\Delta Cyp125A3-\Delta142A2-\Delta125A4$ (pSEVA65g), que no expresa ningún gen heterólogo, acumuló hasta 1 mg/L de progesterona (Fig. R1.12A), siendo la primera vez que se describe la acumulación de este esteroide de forma endógena en *M. smegmatis*, lo que sugiere la existencia de alguna enzima capaz de realizar este paso enzimático en esta bacteria.

Según el tipo de CYP450, Fdx y FdxR, variará la especificidad del correspondiente cofactor redox NAD^+ o $NADP^+$. El *pool* de NADH o de NADPH en la célula bacteriana depende del tipo de fuente de energía que se añada en la biotransformación, siendo el glicerol un sustrato preferente para producir NADH y la

glucosa para producir NADPH. Este es un hecho a tener en cuenta, ya que puede influir finalmente en la actividad del CYP450 (Döhr et al., 2001; Gan et al., 2009; Strohmaier et al., 2019). Para comprobar cuál es el mejor compuesto para utilizar como fuente de energía en la producción de progesterona, se decidió ensayar también la cepa $\Delta Cyp125A3-\Delta 142A2-\Delta 125A4$ (p65_Phsp60_CYP_Fdx_ArR) en las mismas condiciones, pero en presencia de glucosa 2% (p/v). En estas condiciones se produjo 1,62 mg/L de progesterona y 0,13 mg/L de 1,2-dehidroprogesterona (Fig. R1.12), unos valores por debajo de los obtenidos utilizando glicerol. Por lo tanto, se fijó así el glicerol como fuente de energía para las próximas biotransformaciones.

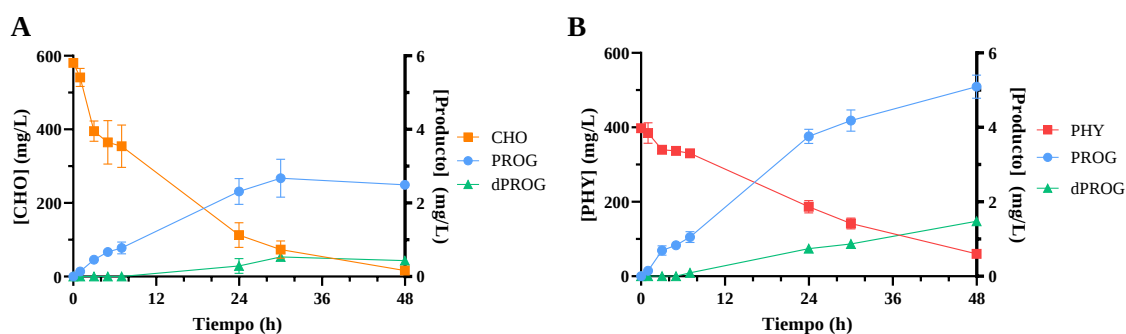


Figura R1.13. Producción de progesterona (PROG) y 1,2-dehidroprogesterona (dPROG) en el tiempo por la cepa $\Delta Cyp125A3-\Delta 142A2-\Delta 125A4$ (p65_Phsp60_CYP_Fdx_ArR), biotransformando 500 mg/L de colesterol (CHO) (A) o de fitoesteroles (PHY) (B). Los valores representan la media de 3 réplicas biológicas.

Una vez determinadas tanto la cepa transformante como la fuente de energía óptima, se decidió ensayar otro tipo de esteroides como sustrato para producir progesterona. Actualmente los fitoesteroides son la principal materia prima utilizada para producir distintos esteroides, ya que es un sustrato barato que se puede conseguir a partir de residuos de agricultura o de industrias como la papelera (Donova y Egorova, 2012; Malaviya y Gomes, 2008). Es por ello que se ensayó la producción en el tiempo de progesterona por parte de la cepa $\Delta Cyp125A3-\Delta 142A2-\Delta 125A4$ (p65_Phsp60_CYP_Fdx_ArR) a partir de 500 mg/L de fitoesteroides de soja, proporcionados por Gadea Biopharma. Como se ha indicado previamente, se añadió glicerol 2% (p/v) como fuente de energía al empezar el experimento y se volvió a añadir a la misma concentración pasadas 24 h. Los resultados se compararon con la producción obtenida utilizando colesterol como sustrato. Como se observa en la Figura R1.13, a partir de 1 h se empieza a detectar progesterona en el medio independientemente del sustrato añadido, y llega a alcanzar una concentración final a las 48 h de 2,5 y de 5,1 mg/L según se utilice colesterol o fitoesteroides, respectivamente. En ambos casos también se observa la acumulación de 1,2-dehidroprogesterona, empezando a detectarse a las 7 h y produciéndose 0,43 y 1,48 mg/L a tiempo final cuando se utilizó colesterol o fitoesteroides, respectivamente. En el caso de la cepa control $\Delta Cyp125A3-\Delta 142A2-\Delta 125A4$ (pSEVA65g), no se detectó en presencia de colesterol la acumulación de progesterona ni

de 1,2-dehidroprogesterona cuando se ensayó la biotransformación de fitoesteroles en *resting cells* (Fig. R1.14).

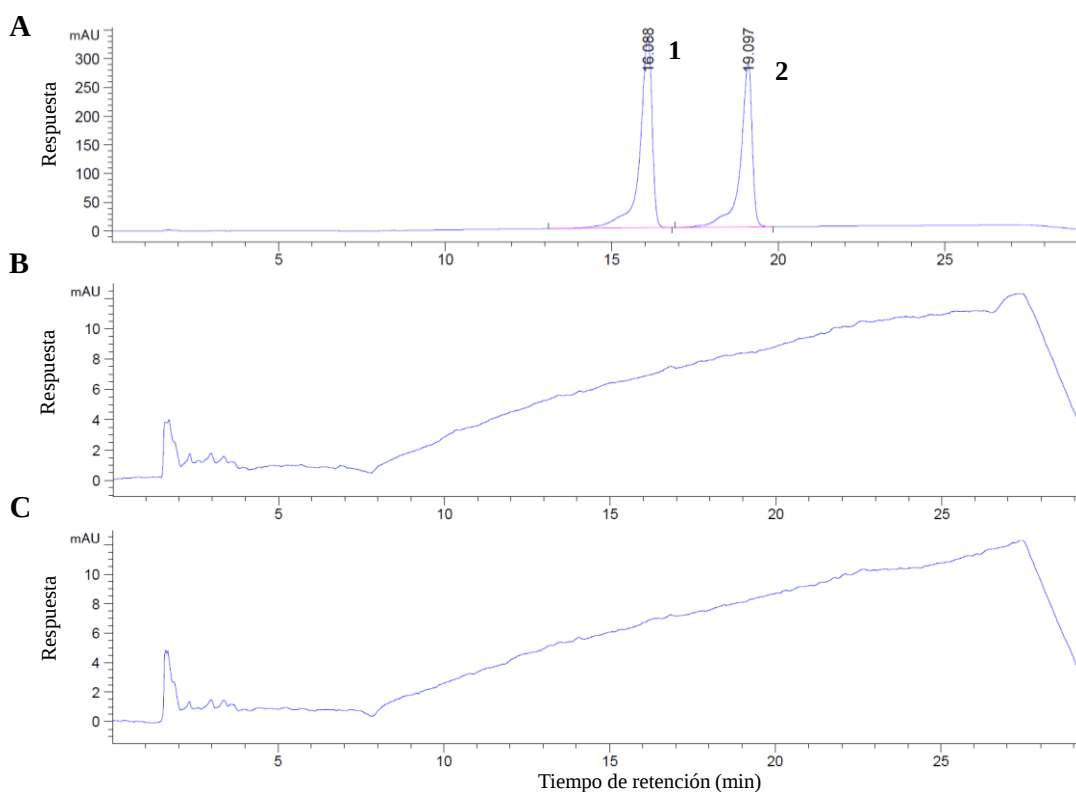


Figura R1.14. Análisis por HPLC de los esteroides producidos por la cepa $\Delta Cyp125A3-\Delta 142A2-\Delta 125A4$ (pSEVA65g). Se compara el cromatograma (A) de los patrones de 1,2-dehidroprogesterona (pico 1) y de progesterona (pico 2) con los cromatogramas obtenidos de la biotransformación de 500 mg/L de fitoesteroles a las 0 h (B) y a las 30 h (C).

El consumo del sustrato de partida presenta diferencias según el origen del esterol. El colesterol desaparece, quedando aproximadamente 16 mg/L a las 48 h habiéndose consumido 564,4 mg/L, mientras que con los fitoesteroles quedan 60 mg/L tras consumir 337,64 mg/L (Fig. R1.13). El rendimiento de esta biotransformación por tanto sería de 19 mg de hormona (progesterona y 1,2-dehidroprogesterona) por g de fitoesteroles, casi 4 veces mayor que los 5 mg de hormona obtenidos por g de colesterol. Respecto a la productividad a las 24 h, a partir de fitoesteroles se obtiene una producción de 0,19 mg/L h⁻¹ del esteroide de interés, mientras que este valor baja hasta 0,11 mg/L h⁻¹ cuando se parte de colesterol. Estos datos confirman que los fitoesteroles son un sustrato adecuado para la producción bacteriana de progesterona, obteniendo incluso unos mayores niveles de esta hormona comparado con el uso de colesterol, lo que posibilita la futura aplicación industrial de este proceso.

1.4.3. Mejora de la producción de progesterona mediante *Golden Standard*

Al inicio de esta Tesis Doctoral no se había descrito todavía la producción de progesterona en *M. smegmatis*. Sin embargo, Karpov et al. (2024) publicaron el año pasado la construcción de una cepa de *M. smegmatis* que expresa el sistema CH/L de mamíferos, compuesto por el CYP450scc bovino junto con la Adx y la AdxR, y capaz de producir pregnenolona desde colesterol. Además, para mejorar la producción de pregnenolona, el hidroxilo del C3 del esterol se bloquea con un grupo metoximetilo evitando modificaciones en el núcleo esteroideo, como la oxidación a un grupo cetona y la isomerización $\Delta^5 \rightarrow \Delta^4$.

Para mejorar la productividad de la biotransformación y conociendo la fuerza de los promotores disponibles para *M. smegmatis* (Fig. R1.5), se ha tratado de aumentar la expresión de las proteínas involucradas en la reacción. En las condiciones ensayadas, el CYP450scc (*Saro_3659*), la Fdx (*Saro_3658*) y la ArR (*Saro_0216*) se expresan con el promotor débil Phsp60 en el plásmido p65_Phsp60_CYP_Fdx_ArR y, para aumentar su transcripción, se eligió el promotor J23100 para los 3 genes, ya que presenta una expresión hasta casi 4 veces mayor (Fig. R1.5). De esta forma, mediante reacciones de GS, se construyó el plásmido p65_J23100_CYP_Fdx_ArR (Tabla M2), en el que cada uno de los 3 genes se expresan de forma individual desde el promotor J23100. Este plásmido se electroporó en el triple mutante $\Delta Cyp125A3\text{-}\Delta I42A2\text{-}\Delta I25A4$ como se indica en el apartado 4.3 de Materiales y métodos. Al comprobar la presencia del plásmido en los transformantes obtenidos se observó sin embargo que presentaban un tamaño mucho menor al esperado (carriles 2-4 de la Figura R1.15A), sugiriendo que algunas secuencias del plásmido se perdían tras su transformación. Las deleciones observadas en estos plásmidos se ratificaron mediante secuenciación por nanoporo en Plasmidsaurus Inc y consistían en la eliminación de parte de los genes que codificaban el CYP450scc y la ArR. La alta expresión de estos genes desde un promotor constitutivo fuerte podría suponer una mayor producción de especies reactivas de oxígeno (ROS) y/o una mayor carga metabólica, ya que la FdxR puede actuar como un sumidero de poder reductor. Por esta razón las bacterias seleccionadas tras la electroporación son las que presentan un plásmido aberrante que no genera toxicidad en la célula.

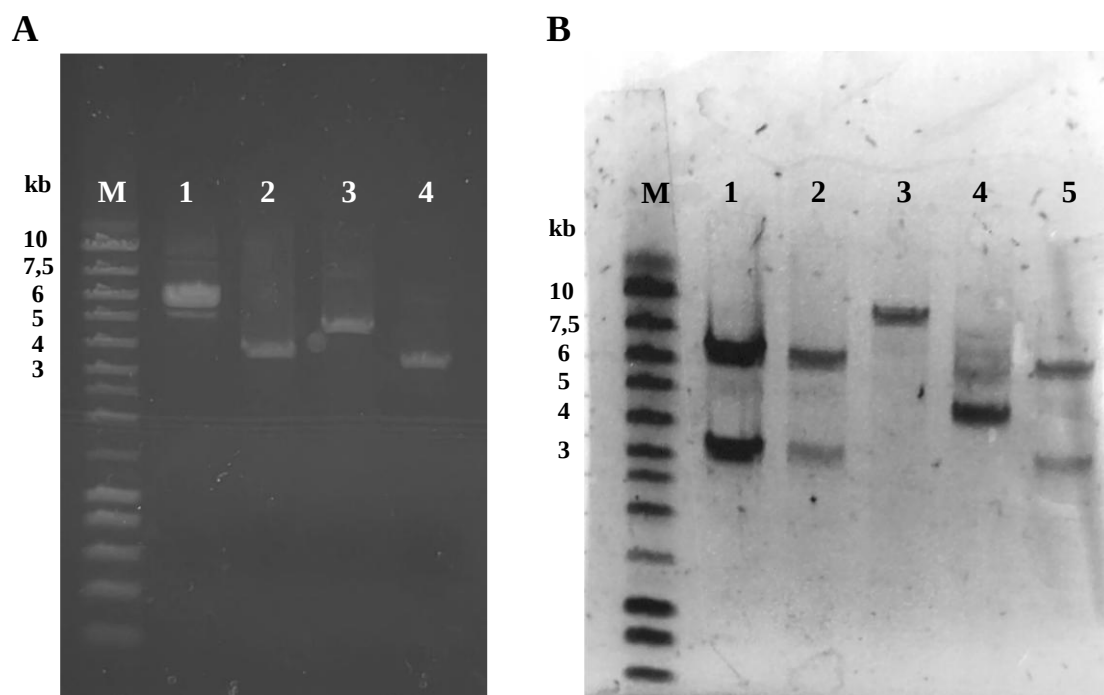


Figura R1.15. Comprobación de los plásmidos extraídos de transformantes de $\Delta Cyp125A3$ - $\Delta I142A2$ - $\Delta I25A4$ por geles de agarosa. (A) Carril M: marcador de peso molecular; carril 1: p65_J23100_CYP_Fdx_ArR; carriles 2-4: plásmidos aberrantes extraídos de 3 clones de $\Delta Cyp125A3$ - $\Delta I142A2$ - $\Delta I25A4$ (p65_J23100_CYP_Fdx_ArR). (B) Carril M: marcador de peso molecular; carril 1: p65_J23100_CYP_Phsp60_Fdx_ArR digerido con *Sph*I; carriles 2-5: plásmidos aberrantes extraídos de 3 clones de $\Delta Cyp125A3$ - $\Delta I142A2$ - $\Delta I25A4$ (p65_J23100_CYP_Phsp60_Fdx_ArR) digeridos con *Sph*I.

Para evitar esto y, al mismo tiempo, mejorar la producción de progesterona, se decidió construir un plásmido en el que la expresión del CYP450scc se controlara con el promotor fuerte J23100, mientras que la expresión de Fdx y ArR se controlara con el promotor constitutivo débil Phsp60. Mediante GS se creó el plásmido p65_J23100_CYP_Phsp60_Fdx_ArR (Tabla M2) que se transformó en la cepa $\Delta Cyp125A3$ - $\Delta I142A2$ - $\Delta I25A4$. Sin embargo, al analizar el tamaño y la secuencia del plásmido que contenían los transformantes obtenidos, se observó que, en algunos casos, aunque su peso molecular era el esperado, presentaban mutaciones puntuales en su secuencia que, o bien modificaban la pauta de lectura, o bien afectaban a la secuencia del RBS (carriles 2 y 5 de la Figura R1.15B) y, en otros casos, existían deleciones de mayor tamaño (carriles 3 y 4 de la Figura R1.15B). Estos resultados sugieren que la sobreexpresión de estas 3 proteínas resulta tóxica para la célula y, por tanto, revelan la necesidad de conseguir un mayor control en su expresión, ya sea por separado o de los 3 genes a la vez. Es por tanto necesario el desarrollo de un sistema de inducción en micobacterias con una mayor potencia, que complemente al sistema TetR/Ptet de expresión más débil, constituyendo el paso a seguir para superar este problema y poder aplicarlo en la mejora de la producción de progesterona.

CAPÍTULO 2. Estudio del metabolismo de la progesterona en *C. tardaugens* NBRC 16725

2.1. Metabolismo de la progesterona en *C. tardaugens*

Publicación 3 (Anexo III)

Título: *Unraveling the key drivers of bacterial progesterone degradation*

Autores: Gabriel Hernández-Fernández, Juan Ibero, José L. García & Beatriz Galán

Revista: *mBio*

Referencia: Hernández-Fernández, G., Ibero, J., García, J.L. & Galán, B. (2025) Unraveling the key drivers of bacterial progesterone degradation. *mBio*, e01077-25. <https://doi.org/10.1128/mbio.01077-25>

Aportación del doctorando: el doctorando Gabriel Hernández ha desarrollado tanto la metodología y la investigación del artículo como el análisis de los datos obtenidos, además de participar en la escritura del borrador del artículo y su revisión.

Resumen:

La progesterona es una de las hormonas sexuales que se puede encontrar como contaminante en suelos y medios acuáticos. A pesar de que distintas especies de microorganismos son capaces de degradar la progesterona, las rutas metabólicas óxicas que llevan a su completa mineralización son todavía desconocidas. En este trabajo se ha investigado el catabolismo de la progesterona en la α -proteobacteria *Caenibius tardaugens* (previamente nombrada como *Novosphingobium tardaugens* ARI-1). Esta bacteria es capaz de degradar una gran variedad de disruptores endocrinos, entre los que se incluyen andrógenos, estrógenos y progestágenos. Se realizó la secuenciación del transcriptoma para comparar la expresión génica en presencia de progesterona y de testosterona como control. El análisis del transcriptoma mostró que existía un aumento de los niveles de expresión de los genes *EGO55_13845* y *EGO55_13860* en presencia de progesterona. Estos genes codifican una Baeyer-Villiger monooxigenasa (BVMO) y una luciferasa-like monooxigenasa (LLM), respectivamente, y se encuentran ubicados junto a un par de genes que codifican proteínas reguladoras, conformando un *cluster* que se denominó *pdc* (*Progesterone degradation cluster*). Estos genes además se encuentran también en otras bacterias capaces de degradar esteroides. Mediante mutagénesis dirigida se construyeron los mutantes simples Δ BVMO y Δ LLM, de los cuales solo Δ LLM presentó una mayor fase *lag* cuando se cultivó con progesterona. Ante una posible redundancia enzimática, se construyó la cepa doble mutante Δ LLM-BVMO que no fue capaz de crecer en progesterona como única fuente de carbono y de energía, confirmando así su papel en el metabolismo de este esteroide. Además, este mutante acumula 1,2-dehidroprogesterona como intermediario. Mediante complementación génica se confirmó que efectivamente las enzimas BVMO y LLM son las responsables de la degradación de progesterona en *C. tardaugens*. Para caracterizar la reacción enzimática catalizada por ambas enzimas, estas se clonaron de forma individual en el plásmido pET y se produjeron en *E. coli* BL21. Los extractos crudos se ensayaron utilizando progesterona como sustrato obteniéndose en ambos casos acetato de testosterona como producto. Así se demostró que estas proteínas, con actividad redundante, catalizan una reacción de Baeyer-Villiger

produciendo acetato de testosterona que, una vez formado se podría degradar a través de la ruta catabólica de la testosterona de esta bacteria. Estos datos sugieren la conexión metabólica entre la degradación de la progesterona y de la testosterona. La BVMO y la LLM se ensayaron con otros sustratos como la pregnenolona, la 1,2-dehidroprogesterona y la 21-hidroxiprogesterona y se observaron diferencias en la actividad específica según el sustrato. La BVMO presenta una mayor actividad frente a la progesterona y a la pregnenolona, mientras que la LLM tiene la misma actividad específica con la progesterona, la pregnenolona y la 1,2-dehidroprogesterona. Esta versatilidad a la hora de transformar distintos sustratos explicaría la importancia biológica de la redundancia enzimática de ambas monooxigenasas.

CAPÍTULO 3. Estudio del catabolismo del colesterol en *C. tardaugens* NBRC 16725

3.1. Análisis del crecimiento de *C. tardaugens* en colesterol

Los esteroides, como el colesterol, son un tipo de esteroides que presentan un grupo hidroxilo en C3 y una larga cadena alifática en C17. Estas moléculas se caracterizan por ser muy apolares, con muy baja solubilidad en agua, y son consideradas compuestos recalcitrantes a la degradación, ya que no son fácilmente metabolizables por los microorganismos y persisten mucho tiempo en el medio ambiente. Entre las bacterias Gram positivas, solo algunas actinobacterias presentan un arsenal de genes encargados del transporte de estos compuestos al interior celular, la degradación de su cadena lateral y la apertura del núcleo esteroideo para su completa degradación. En cuanto a las bacterias Gram negativas, sólo se conocen dos cepas capaces de degradar colesterol: la cepa 72Chol (No clasificada, DSM 12783) y *Sterilobacterium denitrificans*, aunque es en esta última donde se ha estudiado su catabolismo en profundidad (Chiang et al., 2008b; Tarlera y Denner, 2003; Wang et al., 2013b). Sin embargo, en la Tesis Doctoral de Ibero (2022) se describió por primera vez la capacidad de la α -proteobacteria *C. tardaugens* de utilizar el colesterol y otros esteroides con cadena lateral corta como fuente de carbono y de energía en condiciones aerobias (Fig. R3.1). En este estudio se analizó el genoma y se comprobó que *C. tardaugens* no presenta los *clusters* génicos clásicos de degradación de colesterol encontrados en actinobacterias o en *S. denitrificans* (Ibero, 2022). Para determinar la ruta catabólica del colesterol se realizaron análisis transcriptómicos y se construyeron varios mutantes en algunos de los genes que se encontraban inducidos en presencia de colesterol. Comparando la expresión génica de la bacteria en presencia de colesterol frente a testosterona, Ibero (2022) describió un pequeño *cluster*, denominado *cluster cdc* (del inglés *cholesterol degradation cluster*), que consta de un regulador tipo TetR (*EGO55_10195*), un CYP450 (*EGO55_10200*) y una dioxigenasa (*EGO55_10205*) y cuya expresión presentaba una inducción de 1, 34 y 35 veces en presencia de colesterol, respectivamente. Para demostrar el papel de este CYP450 (denominado CYP2113) en la degradación del colesterol se deletionó su gen codificante generando la cepa mutante Δ CYP2113. La cepa no fue capaz de crecer en colesterol como única fuente de carbono y energía y su fenotipo silvestre se restauró cuando se expresó de forma heteróloga el gen mutado. Sin embargo, el crecimiento de la cepa Δ CYP2113 no se vio alterado cuando se utilizó colesteno, un derivado oxidado del colesterol, como fuente de carbono y de energía. Estos resultados demuestran que el gen *EGO55_10200*, que codifica un CYP450, está involucrado en la ruta de degradación del colesterol en *C. tardaugens*, aunque su actividad y su papel específico están aún por determinar.

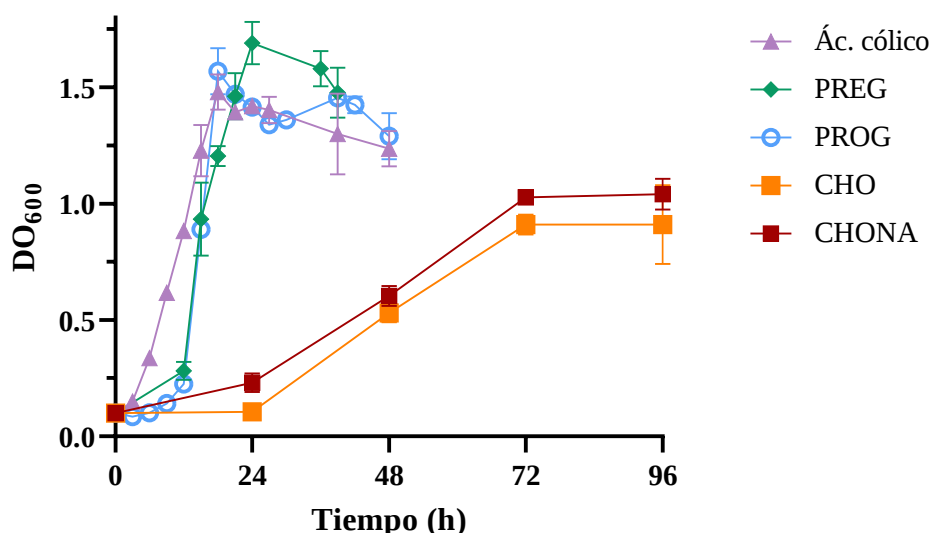


Figura R3.1. Curva de crecimiento de *C. tardaugens* NBRC16725 en diferentes esteroides C21. Ác. cólico: ácido cólico; PREG: pregnenolona; PROG: progesterona; CHO: colesterol; CHONA: colesteno. Los valores representan la media de 3 réplicas biológicas.

A la vista de estos resultados se propone la existencia de un CYP450sc (probablemente el CYP2113) que sería funcionalmente equivalente a los CYP450sc descritos en *N. aromaticivorans* y *N. subterraneum*. Estos CYP450 han sido ensayados tanto *in vitro* (Yamagishi et al., 2011) como *in vivo* (ver Capítulo 1) y son capaces de romper la cadena lateral del colesterol mediante dos monooxigenaciones transformándolo en pregnenolona y liberando isocaproal. Esta pregnenolona finalmente entraría en la ruta de degradación de la progesterona descrita en el Capítulo 2 de esta Tesis Doctoral.

3.2. Actividad del citocromo CYP2113

Para conocer cuál es la reacción enzimática catalizada por CYP2113, el gen *EGO55_10200* se clonó en el plásmido de expresión pET-29a(+), generando el plásmido pET-CYP2113 (Tabla M2), que se transformó en la cepa de *E. coli* BL21 DE3 tal y como está descrito en el apartado 4.1 de Materiales y métodos, obteniendo así la cepa *E. coli* (pET-CYP2113). Esta cepa se cultivó a 37 °C en 50 mL de medio LB hasta la fase exponencial cuando se indujo la expresión con IPTG 0,2 mM. Las células se recogieron tras 3 h de incubación a 30 °C y 200 rpm, se obtuvieron los extractos crudos como se ha descrito previamente comprobándose la sobreproducción de la enzima mediante PAGE-SDS (datos no mostrados). Los ensayos de actividad *in vitro* de este citocromo se realizaron usando estos extractos crudos obtenidos y 0,1 mM de colesterol como sustrato tal y como se describe en el apartado 6.4 de Materiales y métodos. Además, se incorporó a la reacción un sistema heterólogo donador de electrones, conformado por la Fdx y la FdxR de espinaca, y el sistema de regeneración de NADPH basado en la glucosa-6-fosfato deshidrogenasa. La reacción se incubó a RT, con agitación magnética y durante 1 h, tiempo tras el cual la reacción se paró añadiendo acetonitrilo. Los esteroides presentes

se extrajeron con cloroformo y posteriormente se analizaron por GC-MS. Esta técnica nos permite identificar los nuevos picos comparando su patrón de fragmentación con la base de datos de NIST y, al mismo tiempo, cuantificar el colesterol, ya que la columna de HPLC-MS indicada en el apartado 8.2 de Materiales y métodos no resuelve bien este compuesto. Como se puede ver en la Figura R3.2, el CYP2113 no presentó actividad frente a colesterol en las condiciones ensayadas, ya que no se observa ningún pico nuevo correspondiente a algún esteroide al comparar los cromatogramas antes y después de la reacción.

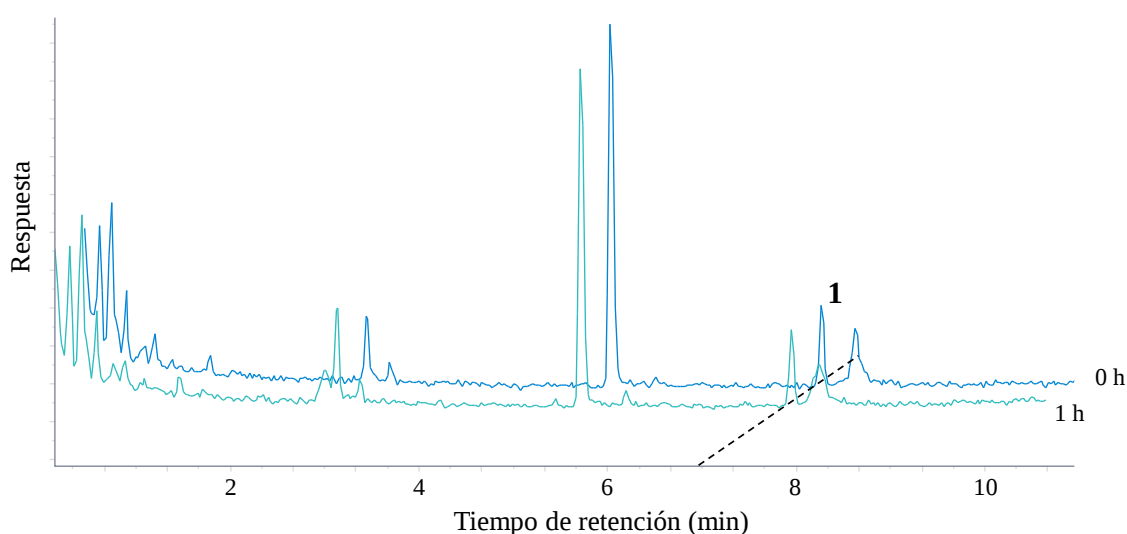


Figura 3.2. Análisis por GC-MS del ensayo enzimático del colesterol con el extracto crudo de *E. coli* (pET-CYP2113) a las 0 h (azul) y 1 h (turquesa). El pico (1) corresponde al colesterol.

Ante este resultado, se decidió probar una aproximación distinta utilizando un par Fdx/FdxR procariota por si el sistema redox eucariota utilizado previamente no es capaz de interactuar con el CYP450 a estudiar. En el genoma de *C. tardeus* no se encontró ninguna Fdx cercana al gen de CYP2113 (*EGO55_10200*) y tampoco se conocen las FdxR que actuarían junto a ellas. Por este motivo, el CYP2113 se sobreexpresó de forma heteróloga junto con la Fdx y la FdxR de *N. aromaticivorans*, utilizadas en el Capítulo 1 de esta Tesis Doctoral para la producción de progesterona. Los genes *Saro_3658* y *Saro_0216* se clonaron formando un operón con el CYP2113 en el plásmido pET-CYP2113, dando lugar a la construcción pET-CYP2113_Fdx_ArR (Tabla M2), que se utilizó a su vez para transformar *E. coli* BL21 DE3. La cepa resultante *E. coli* (pET-CYP2113_Fdx_ArR) se ensayó *in vivo* mediante *resting cells* de igual forma que está descrito en Yamagishi et al. (2011) y como se especifica en el apartado 7.2 de Materiales y métodos. Las bacterias primero se cultivaron a 37 °C en 20 mL de medio mínimo M9 suplementado con casaminoácidos 1% (p/v), ácido 5-aminolevulínico 0,5 mM, FeSO₄ 0,1 mM y glucosa 0,4% (p/v) y se indujo con 0,2 mM de IPTG. Al día siguiente se concentró todo el cultivo en 1 mL de tampón fosfato 50 mM (pH 7,5) con glicerol 2% (p/v), CDX

20 mM y colesterol 0,25 mM disuelto en metanol como sustrato. Se guardó una muestra a las 0 h y otra a las 24 h que posteriormente se extrajeron y se analizaron por GC-MS. Igual que ocurrió con el ensayo *in vitro*, no se observó la aparición de ningún pico de naturaleza esteroidea correspondiente a la biotransformación del colesterol (Fig. R3.3).

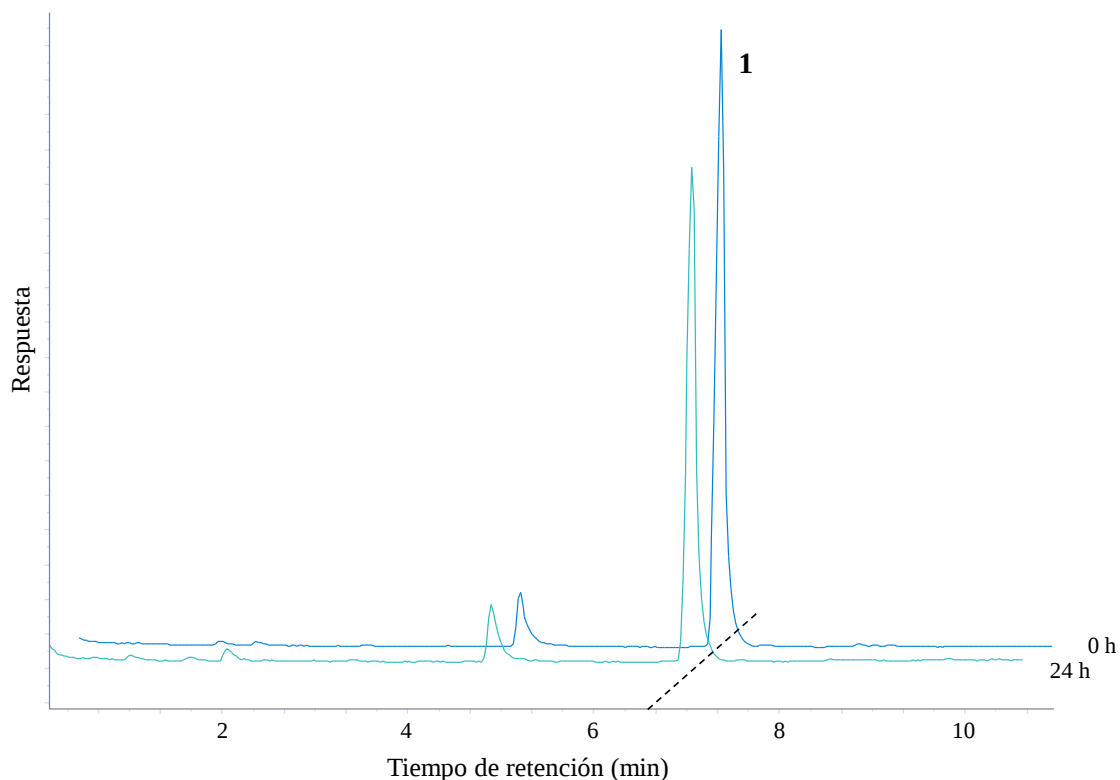


Figura 3.3. Análisis por GC-MS del ensayo de *resting cells* con colesterol con la cepa *E. coli* (pET-CYP2113_Fdx_ArR) a las 0 h (azul) y 24 h (turquesa). El pico (1) corresponde al colesterol.

3.3. Caracterización del mutante Δ LLM-BVMO en colesterol

Debido a que no se ha conseguido demostrar la reacción catalizada por el CYP2113 *in vitro* y con la hipótesis basada en que, como ruta alternativa para la degradación del colesterol, éste se transforma en pregnenolona, se decidió analizar el crecimiento en colesterol de las cepas de *C. tarдаugens* mutantes en las monooxigenasas BVMO y/o LLM (ver Capítulo 2). La delección de ambas monooxigenasas impide el crecimiento de la bacteria en presencia de progesterona o pregnenolona como fuente de carbono y de energía de manera que, si el colesterol es transformado en alguno de estos esteroides C21, el doble mutante Δ LLM-BVMO tampoco sería capaz de crecer en presencia de este esteroide y acumularía el intermediario.

La Figura R3.4 muestra la curva de crecimiento de los mutantes Δ BVMO, Δ LLM, Δ LLM-BVMO y Δ CYP2113 comparado con la cepa silvestre en medio mínimo con 1,33 mM de colesterol como fuente de carbono y de energía. Los mutantes simples Δ BVMO

y Δ LLM mostraron un crecimiento similar al de la cepa silvestre, a pesar de que Δ LLM presentaba una mayor fase *lag* cuando se cultiva en progesterona como se ha descrito previamente. El doble mutante Δ LLM-BVMO, sin embargo, no creció con colesterol en las primeras 96 h, lo que sugiere que el catabolismo de la progesterona está relacionado con el del colesterol en *C. tardaugens*. Sorprendentemente, a partir de las 192 h (8 días) la cepa Δ LLM-BVMO fue capaz de adaptarse y crecer en colesterol (Fig. R3.4).

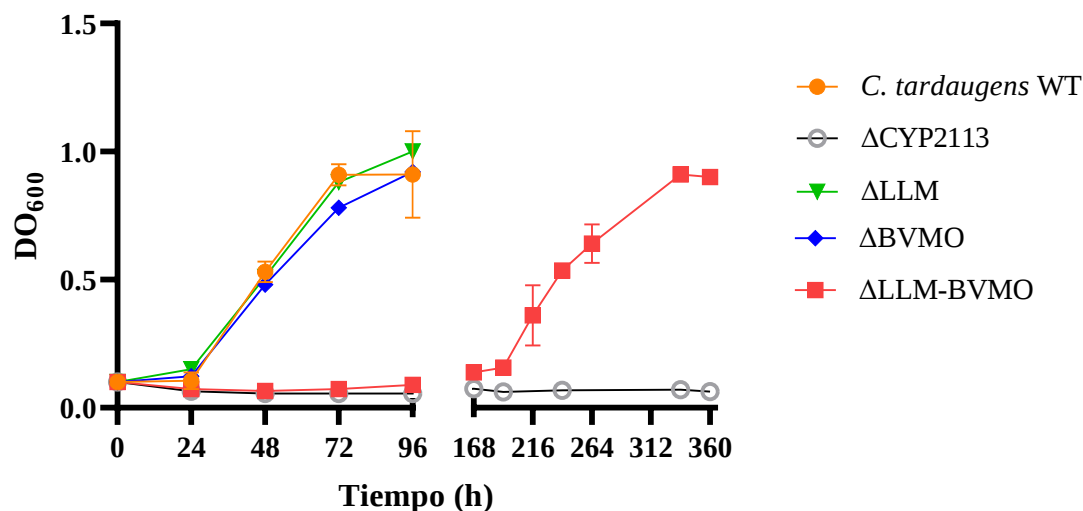


Figura 3.4. Curva de crecimiento de las cepas *C. tardaugens* WT, Δ CYP2113, Δ BVMO, Δ LLM y Δ LLM-BVMO con colesterol como sustrato. Los valores representan la media de 3 réplicas biológicas.

Para comprobar la acumulación de posibles intermediarios se tomaron muestras del medio de cultivo del doble mutante que se analizaron por HPLC. A las 48 h, cuando todavía no se ha observado crecimiento de la cepa, aparece un pico que coincide con el patrón comercial de la 1,2-dehidroprogesterona en el tiempo de retención y en masa ($[M+H^+]=313$), mientras que, a las 240 h, este pico desaparece (Fig. R3.5).

Este resultado demuestra por primera vez que la 1,2-dehidroprogesterona es un metabolito intermediario de la degradación del colesterol en *C. tardaugens*, apoyando la existencia de un citocromo similar al CYP450scc (posiblemente el CYP2113) que cortaría la cadena lateral del colesterol para formar pregnenolona que, a su vez, se transformaría en progesterona por acción de una 3β -HSD (Fig. R1.8). La progesterona a continuación se puede deshidrogenar por una de las KstD codificadas en el genoma de *C. tardaugens* originando 1,2-dehidroprogesterona.

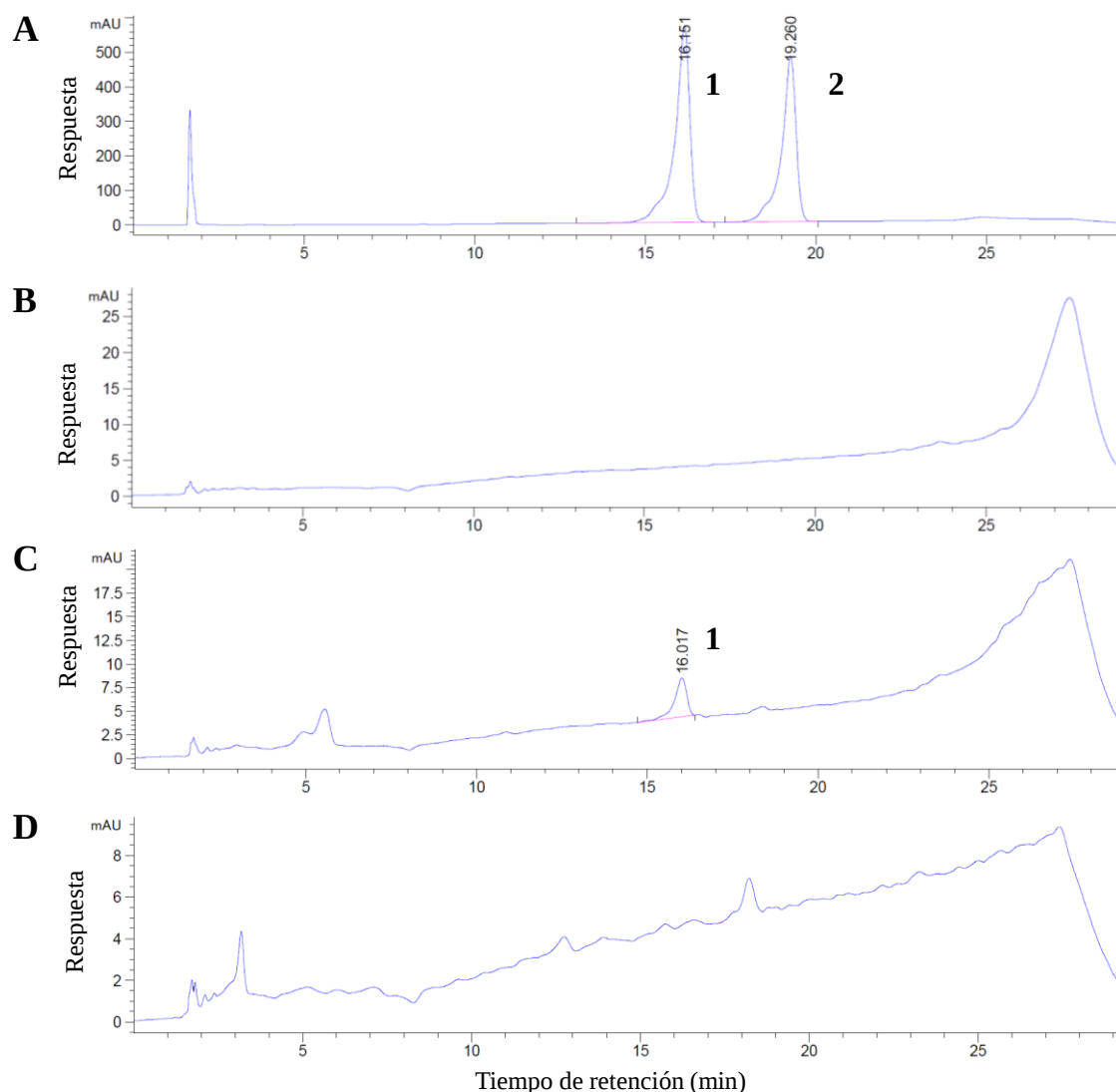


Figura R3.5. Análisis por HPLC de los intermediarios acumulados en el medio de cultivo de Δ LLM-BVMO en colesterol 1,33 mM. Se comparó el cromatograma de los patrones (A) de 1,2-dehidroprogesterona (pico 1) y progesterona (pico 2) con los cromatogramas obtenidos a las 0 h (B), 48 h (C) y 240 h (D) del crecimiento.

La acción del CYP450scc sobre el colesterol libera pregnenolona y el aldehído isocaproal como productos al producirse el corte de la cadena lateral (Strushkevich et al., 2011). El crecimiento de la cepa Δ LLM-BVMO en colesterol pasadas las 192 h sugiere que estos compuestos acaban siendo metabolizados por la bacteria, probablemente tras un periodo de adaptación, e incluso se llega a alcanzar la misma DO_{600} final que con la cepa silvestre (Fig. 3.4). Este doble mutante no debería ser capaz de utilizar la pregnenolona o la progesterona como fuente de carbono y de energía, pero como se puede observar en la Figura R3.5C, la 1,2-dehidroprogesterona podría ser mineralizada, o al menos transformada en otro compuesto, ya que desaparece cuando la cepa empieza a crecer finalmente en colesterol.

3.4. Aplicación de la cepa Δ LLM-BVMO para producir progesterona a partir de colesterol

De igual forma que *M. smegmatis* se puede utilizar como factoría celular para la producción de progesterona a partir de esteroides, el doble mutante Δ LLM-BVMO de *C. talarum* sería un buen candidato para la obtención de esta hormona biobasada sin depender de reacciones químicas. Por ello se ensayó la conversión de colesterol en progesterona en esta cepa mediante *resting cells* utilizando colesterol como materia prima.

Las bacterias se cultivaron primero en medio rico NB hasta la fase estacionaria complementando el medio con ácido 5-aminolevulínico 1 mM y FeSO_4 1 mM. Después se concentraron a una $\text{DO}_{600}=16-18$ en medio NB con 500 mg/L de colesterol como sustrato disuelto en CDX. En este caso se usó el medio NB como fuente de energía ya que *C. talarum* no es capaz de metabolizar la glucosa ni el glicerol (Fujii et al., 2003). Se fueron tomando muestras durante 30 h de las que se extrajeron los esteroides presentes y se analizaron por HPLC y GC-MS. En ninguno de los tiempos analizados se detectó progesterona. Sin embargo, se detectó 1,2-dehidroprogesterona desde la primera hora, alcanzándose una concentración máxima de 0,31 mg/L a las 7 h de incubación para, posteriormente, decaer hasta los 0,23 mg/L a las 48 h (Fig. R3.6). Se observó que aproximadamente la mitad del colesterol (232 mg/L) queda sin biotransformar a las 48 h del ensayo, habiéndose consumido solo 264,3 mg/L. La productividad del proceso a las 7 h es 0,044 mg/L h⁻¹ y baja hasta los 0,01 mg/L h⁻¹ a las 24 h. En cuanto al rendimiento final, es de 0,9 mg de producto por 1 g de colesterol.

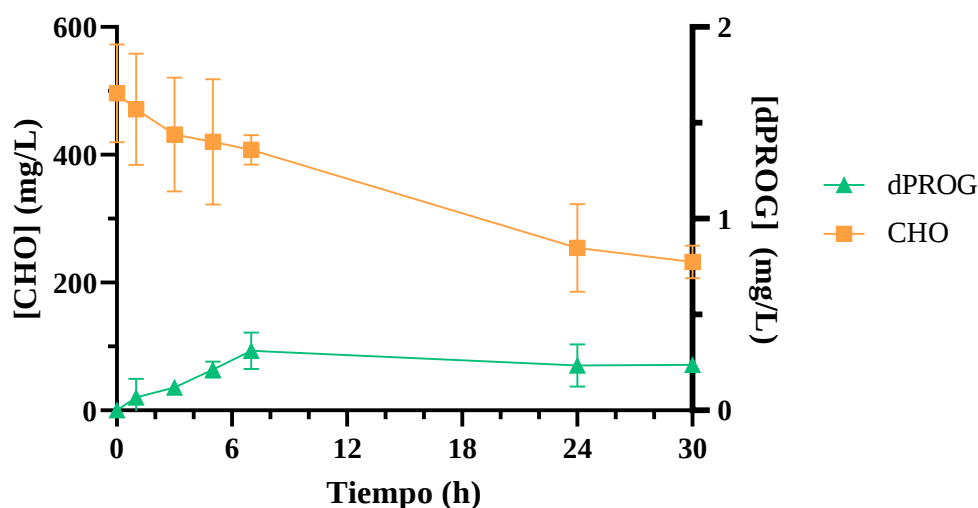


Figura R3.6. Producción de 1,2-dehidroprogesterona (dPROG) en el tiempo a partir de 500 mg/L de colesterol (CHO) en la cepa Δ LLM-BVMO mediante ensayo en *resting cells* con medio rico NB. Los valores representan la media de 3 réplicas biológicas.

3.5. Análisis *in silico* de CYP450, Fdx y FdxR inducidos en colesterol

Por lo general, el plegamiento y la actividad de los citocromos dependen de grupo hemo y de una correcta interacción entre el citocromo, la ferredoxina y la ferredoxina reductasa para que se puedan realizar los ciclos de oxidación-reducción necesarios. El gen del CYP2113 (*EGO55_10200*) no presenta una ferredoxina cercana en el genoma, por lo que se desconoce cuál podría ser el par Fdx/FdxR necesario para su actividad. Con el fin de identificar genes que codifican Fdx, FdxR u otros genes con una posible implicación en el metabolismo del colesterol y, conociendo los genes implicados en la degradación de la progesterona, se han reexaminado los datos transcriptómicos obtenidos por Ibero (2022) en colesterol frente a testosterona y se han comparado con el análisis transcriptómico de *C. tarдаugens* en progesterona frente a testosterona (ver Capítulo 2).

En el genoma de *C. tarдаugens* podemos encontrar 20 genes que codifican CYP450, de los cuales 14 de ellos presentan un $FC \geq 1$ en presencia de colesterol (Tabla R3.1), y únicamente hay 8 genes que codifican Fdx, lo que sugiere que tiene que existir una versatilidad entre las Fdx de manera que algunas de ellas sean capaces de actuar como parejas redox de varios CYP450 o de monooxigenasas. Solo los pares *EGO55_03165-EGO55_03170*, *EGO55_03275-EGO55_03280*, *EGO55_04870-EGO55_04875* y *EGO55_05540-EGO55_05545* codifican un CYP450 con una Fdx cercana, lo que apunta a su posible actuación conjunta para catalizar reacciones redox, y únicamente *EGO55_03275-EGO55_03280* ($FC=2$ y 11) y *EGO55_04870-EGO55_04875* ($FC=5$ y 17) presentan un $FC \geq 2$ en presencia de colesterol. Estos niveles de inducción son inferiores a los observados en el CYP2113 ($FC=34$), sin embargo, no se encontró ninguna Fdx a su alrededor en el genoma. De todas ellas, los genes *EGO55_05120* y *EGO55_19875* codifican Fdx con un FC de 2 y 6, respectivamente, y sus transcritos por millón (TPM) rondan entre 43 y 621 (Tabla R3.1). Las 2 Fdx restantes (*EGO55_00615* y *EGO55_13190*) están reprimidas en presencia de colesterol ($FC=-1$ y -2 , respectivamente), pero sus TPMs se encuentran entre 330 y 6595, lo que indica una expresión basal de estos genes puede ser suficiente para formar parte del sistema redox del CYP2113 u otros CYP450 de esta bacteria. Ninguno de estos genes se encuentra diferencialmente expresado ($|FC| \geq 2$; $FDR < 0,05$) en presencia de progesterona frente a testosterona.

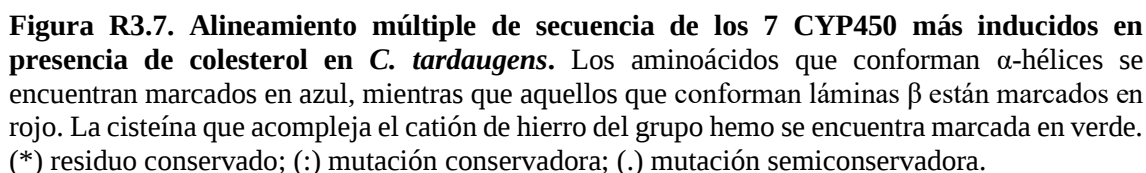
Tabla R3.1. Niveles de expresión de genes que codifican ferredoxinas, ferredoxinas reductasas o citocromos P450. TPM: transcritos por millón; CHO: colesterol; TES: testosterona; PROG: progesterona; FC: *fold change*; CYP450: citocromo P450. Los valores de TPM corresponden a la media de 3 réplicas biológicas analizadas.

Gen	Función	TPMs CHO	TPMs TES	FC CHO vs TES	TPMs PROG	TPMs TES	FC PROG vs TES
<i>EGO55_00615</i>	Ferredoxina	330	355	-1	45	41	1

EGO55_03005	CYP450	718	115	6	6	7	-1
EGO55_03010	CYP450	429	125	3	5	5	-1
EGO55_03165	Ferredoxina	88	131	-1	19	18	-1
EGO55_03170	CYP450	638	324	2	15	14	1
EGO55_03275	CYP450	422	188	2	11	13	-1
EGO55_03280	Ferredoxina	40	4	11	11	18	-1
EGO55_04870	CYP450	370	82	5	9	9	-1
EGO55_04875	Ferredoxina	71	4	17	6	6	-1
EGO55_05120	Ferredoxina	43	8	6	112	132	-1
EGO55_05540	CYP450	381	257	1	23	19	1
EGO55_05545	Ferredoxina	105	62	2	28	30	-1
EGO55_05555	CYP450	606	406	1	15	13	1
EGO55_05560	CYP450	712	344	2	11	15	-1
EGO55_06005	CYP450	269	217	1	6	7	-1
EGO55_09415	Ferredoxina reductasa	4997	12403	-2	361	353	1
EGO55_10200	CYP450	19740	577	34	14	16	-1
EGO55_10220	CYP450	453	166	3	3	4	-1
EGO55_13190	Ferredoxina	6595	10021	-2	782	711	1
EGO55_14220	CYP450	639	200	3	9	9	-1
EGO55_14410	CYP450	635	278	3	11	14	-1
EGO55_18145	CYP450	586	335	2	17	21	-1
EGO55_19875	Ferredoxina	621	125	5	18	23	-1

Ibero (2022) propone los genes *EGO55_01285*, *EGO55_05820*, *EGO55_06710* y *EGO55_09415* como posibles FdxR en *C. tardeaugens* según su anotación en el genoma. No obstante, mediante un alineamiento de proteínas con la herramienta BLAST, únicamente la proteína *EGO55_09415* presenta una identidad del 74% y del 37% frente a las FdxR de *N. aromaticivorans* (ArR) y de *P. putida* (CamA), respectivamente, lo que indica que *EGO55_09415* podría ser la FdxR de *C. tardeaugens*. Aunque su expresión en presencia de colesterol se encuentra ligeramente reprimida con un FC=-2, sus TPMs se aproximan a 5000 (Tabla R3.1), lo que puede constituir unos niveles de expresión basales suficientes para actuar junto con el resto de las Fdx y los CYP450. La expresión de *EGO55_09415* en presencia de progesterona frente a testosterona es similar en ambas condiciones (FC=1).

Aunque la mutación del CYP2113 impide el crecimiento de la bacteria en colesterol, la cepa pudo crecer con colesteno como única fuente de carbono y de energía, lo que sugiere que la colesteno se metaboliza al menos con otro de los CYP450 restantes. En este sentido, cuando se analizó la estructura primaria de 7 de los CYP450 más inducidos (*EGO55_03005*, *EGO55_03010*, *EGO55_04870*, *EGO55_10200*, *EGO55_10220*, *EGO55_14410*, *EGO55_14220*; FC=3-34) mediante un alineamiento múltiple con CLUSTALW, se observó que se conservan muy pocos aminoácidos entre las secuencias primarias, como por ejemplo la cisteína que acompleja el átomo de hierro del grupo hemo (Fig. R3.7). Sin embargo, cuando se utilizó AlphaFold2 y el software



118

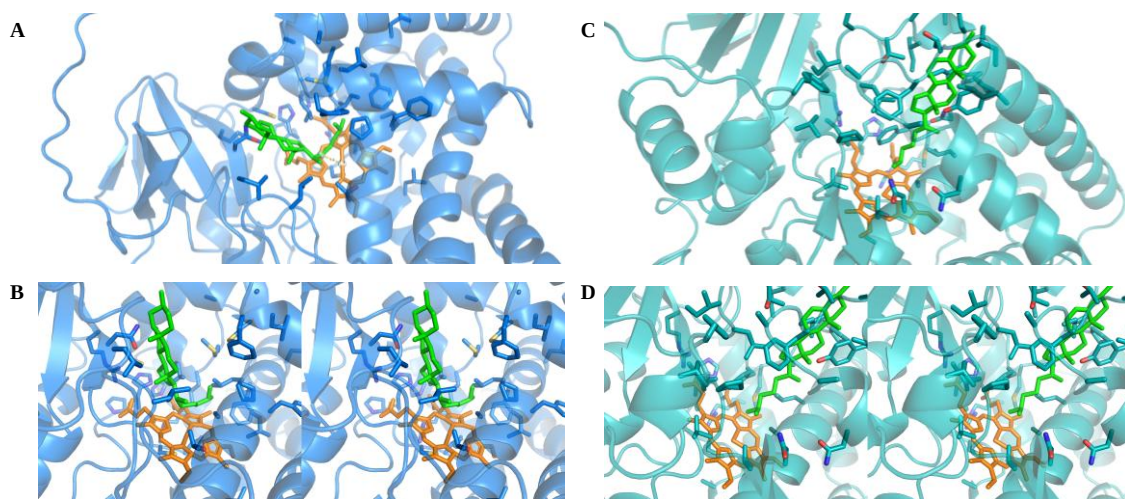


Figura R3.8. Predicción de la estructura de los CYP450 EGO55_10220 (A y B) y EGO55_03005 (C y D) y su interacción con el grupo hemo y el colesterol como ligandos. El grupo hemo está pintado en color naranja, mientras que el colesterol está en verde. Los paneles B y D constituyen la imagen en *wall eyed stereo*.

3.6. Análisis *in silico* de otros genes inducidos en colesterol

El *cluster* *pdc* en presencia de colesterol presentó un FC de entre 1 y 6 y se observó una inducción de la expresión de BVMO y LLM (FC=2 y FC=6, respectivamente) (Tabla R3.2), lo que concuerda con la hipótesis de que el colesterol se transforma en pregnenolona/progesterona que a su vez induciría la expresión del *cluster* *pdc* para su mineralización en *C. tardaugens*. Analizando en detalle los genes colindantes al *cluster* *pdc*, se encontró una región comprendida entre los genes *EGO55_13865* y *EGO55_13910* con una elevada inducción en presencia de colesterol y que podrían constituir una expansión del *cluster* *pdc*. En esta región se localizan genes con un FC entre -2 y 57 que codifican cuatro oxidorreductasas (*EGO55_13865*, *EGO55_13875*, *EGO55_13880* y *EGO55_13890*), una alfa/beta hidrolasa (*EGO55_13870*), una proteína hipotética (*EGO55_13885*), un receptor TonB (*EGO55_13895*), una epimerasa/deshidratasa (*EGO55_13900*), una CoA transferasa (*EGO55_13905*), una enoil-CoA hidratasa (*EGO55_13910*), una proteína de la familia de LysR (*EGO55_13915*), dos acil-CoA deshidrogenasas (*EGO55_13920* y *EGO55_13925*) y un regulador de la familia de TetR (*EGO55_13930*) (Tabla R3.2). El gen *EGO55_13890* que codifica una oxidorreductasa es el más inducido y presenta similitud con una ciclohexanona monooxigenasa de *N. arvoryzae*, una familia de proteínas que catalizan una monooxigenación de Baeyer-Villiger sobre una amplia variedad de sustratos, como puede ser cetonas o aldehídos para transformarlos en ésteres, lactonas o epóxidos (Romero et al., 2016). Los receptores TonB, como el codificado en *EGO55_13895*, se han propuesto como posibles responsables en el transporte de moléculas apolares (Ibero, 2022; Lin et al., 2015), y su inducción en presencia de colesterol (FC=14) podría deberse a su implicación en este proceso. El resto de enzimas que deshidrogenan o hidratan derivados de CoA podrían tener un papel en la degradación de alquenos como puede ser el subproducto isocaproal obtenido de la rotura de la cadena lateral del colesterol. Cuando se analiza el perfil

transcriptómico de *C. taradogens* en presencia de progesterona frente a testosterona, únicamente el *cluster* *pdc* se encuentra inducido con un FC entre 3 y 139, mientras que la región anexa *EGO55_13865-EGO55_13930* presenta un FC entre -1 y 1, sugiriendo que sus niveles de expresión no cambian en respuesta a progesterona.

Tabla R3.2. Niveles de expresión de la región *EGO55_13845-13930*. TPM: transcritos por millón; CHO: colesterol; TES: testosterona; PROG: progesterona; FC: *Fold change*; CYP450: citocromo P450. Los valores de TPM corresponden a la media de 3 réplicas biológicas analizadas.

Gen	Función	TPMs CHO	TPMs TES	FC CHO vs TES	TPMs PROG	TPMs TES	FC PROG vs TES
<i>EGO55_13845</i>	Baeyer-Villiger monooxigenasa	2238	1210	2	4331	28	139
<i>EGO55_13850</i>	Regulador transcripcional de la familia TetR/AcR	2439	1188	2	25	38	-1
<i>EGO55_13855</i>	Regulador transcripcional HTH	896	862	1	112	53	3
<i>EGO55_13860</i>	Luciferasa-like monooxigenasa	12990	2217	6	1899	60	32
<i>EGO55_13865</i>	Oxidorreductasa dependiente de NADP	2539	1216	2	28	30	-1
<i>EGO55_13870</i>	Alfa/beta hidrolasa	3423	479	7	14	15	-1
<i>EGO55_13875</i>	Oxidorreductasa SDR dependiente de NAD(P)	4245	259	16	16	15	1
<i>EGO55_13880</i>	Oxidorreductasa SDR	2900	405	7	24	23	1
<i>EGO55_13885</i>	Proteína hipotética	1704	3061	-2	115	147	-1
<i>EGO55_13890</i>	Oxidorreductasa dependiente de FAD/NAD(P)	87895	1554	57	16	20	-1
<i>EGO55_13895</i>	Receptor TonB	2557	189	14	4	6	-1
<i>EGO55_13900</i>	Epimerasa/deshidratasa	757	541	1	10	13	-1
<i>EGO55_13905</i>	CoA transferasa	377	366	1	14	14	1
<i>EGO55_13910</i>	Crotonasa/enoil-CoA hidratasa	1247	801	2	28	28	-1
<i>EGO55_13915</i>	Proteína de unión a sustrato de la familia de LysR	1034	837	1	39	43	-1
<i>EGO55_13920</i>	Acil-CoA deshidrogenasa	605	655	-1	25	25	1
<i>EGO55_13925</i>	Acil-CoA deshidrogenasa	380	586	-2	26	24	1
<i>EGO55_13930</i>	Regulador transcripcional de la familia TetR/AcR	1619	2785	-2	162	149	1

Esta región se encuentra conservada en mayor o menor medida con respecto a otros *Novosphingobium* o bacterias degradadoras de esteroides como *C. estronivorus* (Fig. R3.9). Genes con una alta similitud en su secuencia de aminoácidos a la oxidorreductasa EGO55_13890 se encuentran presentes en todos ellos y, en el caso de *C. estronivorus*, también se encuentra próximo a su *cluster* *pdc* homólogo, junto con genes similares a la alfa/beta hidrolasa EGO55_13870 y la oxidorreductasa EGO55_13865. Estos 3 genes, junto con el del receptor TonB EGO55_13890 y el del regulador TetR EGO55_13930, se encuentran también en *N. arvoryzae*. Cuando se compara esta región con las especies *N. aromaticivorans* y *N. subterraneum*, en los que se ha estudiado la actividad CH/L de sus CYP450scc, solo se encuentran presentes genes similares a la oxidorreductasa EGO55_13865, a la CoA transferasa EGO55_13905, a la enoil-CoA hidratasa EGO55_13910 y a la acil-CoA deshidrogenasa EGO55_13920, todas ellas formando también un pequeño *cluster*. Cabe destacar que en *N. aromaticivorans*, el gen de la oxidorreductasa EGO55_13890 está a solo 3,5 kb de la BVMO homóloga a EGO55_13845 de *C. tardaugens*.

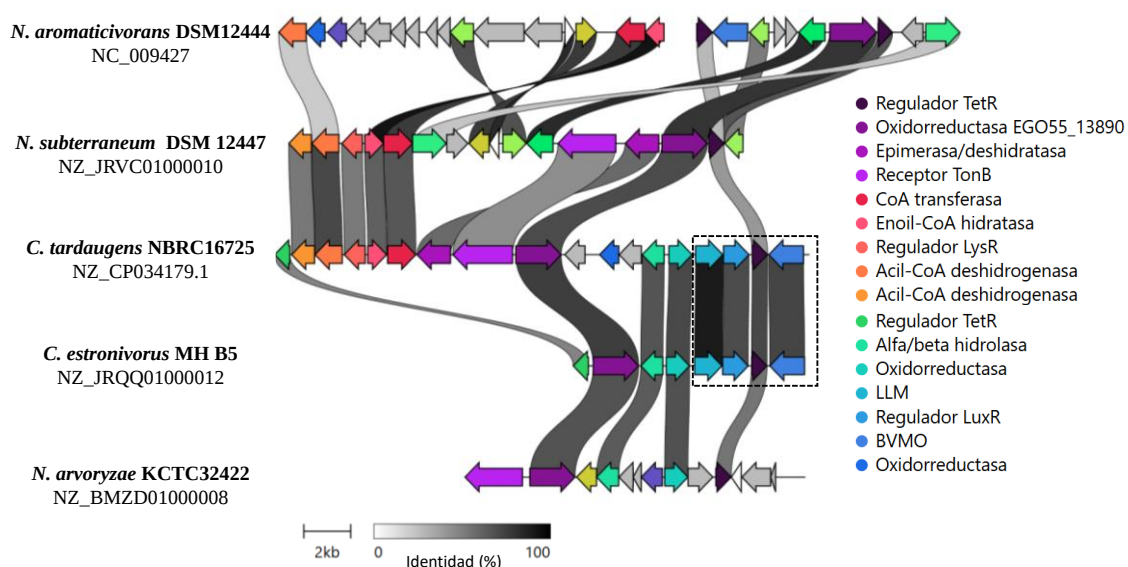


Figura R3.9. Organización génica de la región EGO55_13845-13930 en *C. tardaugens* comparado con otras bacterias relacionadas. El *cluster* *pdc* se encuentra marcado en un cuadro.

Estos resultados sugieren que hay que continuar con un estudio funcional en profundidad tanto de la organización génica como de los procesos bioquímicos para entender cuál es el papel de cada enzima en la degradación, no solo del colesterol, sino también de sus productos pregnenolona e isocaproal. Conocer los complejos CYP450scc/Fdx/FdxR encargados de catalizar la ruptura de la cadena lateral del colesterol y de la colesteno en *C. tardaugens* arrojaría luz sobre un nuevo metabolismo de estos esteroides en bacterias Gram negativas y, además, permitirían una futura aplicación de estas enzimas en procesos biotecnológicos. Por otro lado, estudiar el

catabolismo del isocaproal conllevaría la descripción por primera vez de la degradación de este compuesto en bacterias.



VII. DISCUSIÓN

1. Utilización de *M. smegmatis* como factoría celular para la producción de esteroides con interés industrial

1.1. Producción de esteroides C19 y C22

El AD y el 4-HBC son dos esteroides C19 y C22, respectivamente, que se pueden utilizar como punto de partida para la síntesis de distintas moléculas esteroideas con interés farmacéutico, como puede ser la testosterona, la progesterona o los corticosteroides (Donova y Egorova, 2012; Pan et al., 2023; Peng et al., 2021). Ambas sintonas se obtienen principalmente en micobacterias modificando y delecionando genes involucrados en el metabolismo de los esteroides. En esta Tesis Doctoral, se ha utilizado *M. smegmatis* como modelo porque en este microorganismo el metabolismo y el transporte del colesterol se han estudiado en profundidad (Galán et al., 2017; García-Fernández et al., 2017b, 2015; Uhía et al., 2011a; Uhía et al., 2011b).

El mutante MS6039-5941 de *M. smegmatis* tiene delecionadas los genes que codifican las enzimas KstD y KshB y por tanto produce AD a partir de esteroides, pero además acumula cierta cantidad de 4-HBC como subproducto (Galán et al., 2017). Del mismo modo, cuando se interrumpe la ruta de degradación para producir 4-HBC como producto principal en micobacterias, se ha observado que el esteroide AD también se acumula como subproducto (Han et al., 2024; Liu et al., 2023b; Peng et al., 2021). La presencia de dos subproductos de estructura similar no solo disminuye el rendimiento del proceso, sino que también dificulta la purificación de los compuestos, aumentando el coste de producción. En este trabajo nos hemos centrado en identificar y actuar sobre las enzimas claves responsables del metabolismo del 4-HBC para optimizar *M. smegmatis* como factoría celular con el objetivo de producir sintonas C19 o C22.

Aunque en el metabolismo aeróbico bacteriano de los esteroides se sabía que el intermediario 22-OH-BCN-CoA se transforma en 3-OPA mediante una reacción retroaldol y después se reduce hasta formar 4-HBC (Fig. I4), al inicio de esta Tesis Doctoral, las enzimas responsables de esta catálisis no se encontraban descritas en ninguna de las actinobacterias estudiadas. En esta Tesis se ha identificado un *cluster* génico, denominado 4-HBC, que codifica la aldolasa msSal, la reductasa msRed, la hidratasa msShy y la deshidrogenasa msACAD. Todas estas enzimas tienen una alta similitud a las enzimas implicadas en la degradación de los ácidos biliares.

La aldolasa msSal se delecionó en la cepa MS6039-5941 productora de AD para mejorar su producción, y se obtuvo la cepa MS6039-5941-6561 que no acumuló 4-HBC como subproducto. Este mutante aumentó un 60% la producción de AD a partir de colesterol porque el flujo metabólico se canaliza mayoritariamente hacia su obtención, ya que la producción de 4-HBC se encuentra bloqueada. Esta misma respuesta la observaron posteriormente Zhao et al. (2023) en *M. neoaurum*, donde también se identificó el *cluster* HBC. Cuando delecionaron la aldolasa Atf1, homóloga a msSal, las cepas resultantes no acumulaban ningún esteroide C22 y también aumentaban la producción de AD y ADD en sus respectivos mutantes en un 25% y 13%, respectivamente. Todo ello confirma que esta aldolasa está involucrada en el metabolismo del 4-HBC y que su función no es

sustituida por otra enzima redundante. Este mismo *cluster* se encuentra también en *M. fortuitum*, sin embargo, el estudio de su implicación en el metabolismo del 4-HBC no se ha realizado ni aplicado todavía a mutantes productores de esteroides C19 o C22 (Han et al., 2024; Liu et al., 2023a, 2023b). Zhao et al. (2023) describieron además un regulador denominado KstR3 presente en el *cluster* HBC que actúa reprimiendo la expresión de los genes presentes en el *cluster*. En el caso de *M. smegmatis*, este regulador está presente en el *cluster* 4-HBC codificado por el gen *MSMEG_6564*, sin embargo, su probable función represora, su secuencia operadora y el inductor o inductores todavía no están descritos.

Considerando que la delección de la aldolasa msSal es clave para mejorar la producción de esteroides C19, se consideró que su sobreexpresión en una cepa productora de 4-HBC aumentaría su acumulación. Aunque se habían descrito varias cepas de *M. neoaurum* y *M. fortuitum* capaces de acumular 4-HBC cuando se elimina la actividad β -hidroxiacil-CoA deshidrogenasa Hsd4A (Han et al., 2024; Liu et al., 2023b; Peng et al., 2021; Xu et al., 2016), encargada de oxidar el 22-OH-BCN-CoA a 22-OBNC-CoA (Fig. I4), esto no se había descrito en *M. smegmatis*. En esta Tesis Doctoral se delecionó el gen *hsd4A* (*MSMEG_5903*) de *M. smegmatis* sobre la cepa mutante MS6039-5941 productora de AD, obteniendo así el triple mutante MS6039-5941-5903. Esta nueva cepa produce hasta 5,6 g/L de 4-HBC a partir de 14 g/L de fitoesteroides, un valor similar al observado en *M. fortuitum* ATCC 35855, donde la producción de 4-HBC por la cepa MFKD Δ kshA Δ hsd4A alcanza 5,29 g/L a partir de 10 g/L de fitoesteroides (Liu et al., 2023a), mientras que en *M. neoaurum* la cepa XII Δ hsd4A Δ kstD123 produce alrededor de 15 g/L de 4-HBC a partir de 40 g/L de fitoesteroides (Xu et al., 2016). Sin embargo, el mutante MS6039-5941-5903 acumula el esteroide C19 AD como subproducto, lo que se ha observado también en el resto de micobacterias productoras de 4-HBC, sugiriendo la existencia de redundancia enzimática en este paso del metabolismo del colesterol (Han et al., 2024; Liu et al., 2023b; Peng et al., 2021; Xu et al., 2016).

Con el objetivo de redirigir totalmente el flujo del metabolismo del colesterol hacia la producción de 4-HBC y conociendo la importancia de la aldolasa msSal en el proceso, se decidió sobreexpresar esta enzima en la cepa mutante MS6039-5941-5903. Contrariamente a lo esperado, esta aproximación no supuso ninguna mejora en la producción de 4-HBC observándose una disminución del crecimiento de la nueva cepa en presencia de colesterol, probablemente debido a la acumulación del intermediario aldehído 3-OPA. Para evitar la acumulación tóxica de este aldehído, recuperar el crecimiento normal de la cepa y conseguir una mejora en la producción de 4-HBC, se decidió sobreexpresar también la reductasa msRed o la reductasa msOpccR de *M. smegmatis* para reducir el 3-OPA a 4-HBC (Fig. I4). La actividad de la reductasa msRed no estaba descrita hasta este momento, pero su gen codificante se encuentra en el *cluster* 4-HBC junto con la aldolasa msSal, sugiriendo su relación con el metabolismo del 4-HBC. Curiosamente, esta reductasa no se encuentra en el *cluster* HBC en el genoma de *M. neoaurum*, pero sí se encuentra en el genoma de *M. fortuitum* (Zhao et al., 2023). La actividad de esta reductasa estaría desempeñada en *M. neoaurum* por la acción de la reductasa mnOpccR, cuyo papel se ha descrito en la reducción del 3-oxo-4-pregnen-20-carboxil-CoA (3-OPC-CoA) y del 3-OPA al 4-HBC (Peng et al., 2021). Peng et al. (2021)

estudiaron la actividad de msOpccR de *M. smegmatis*, la reductasa homóloga a mnOpccR de *M. neoaurum*, y se comprobó que es capaz de reducir el 3-OPA a 4-HBC. Cuando las reductasas msRed o msOpccR se sobreexpresaron por separado junto con la aldolasa msSal en el mutante MS6039-5941-5903, no solo se recuperó su crecimiento normal, sino que aumentó la producción de 4-HBC, siendo la primera vez que se mejora el rendimiento de este proceso utilizando la aldolasa y la reductasa involucradas en este metabolismo. Las nuevas cepas aumentaron su producción en un 127% y 61% cuando se ensayaron con fitoesteres al expresar la reductasa msRed o la reductasa msOpccR, respectivamente, y llegando a una producción de 12,7 y 9,0 g/L desde 14,0 g/L de fitoesteres, respectivamente. En otras micobacterias, la estrategia utilizada para mejorar la producción de 4-HBC en *M. fortuitum* consistió en mutar la tiolasa FadA5, de manera que el mutante MFKD Δ kshA Δ hsd4A Δ fadA5 acumuló 5,4 g/L de 4-HBC a partir de 10,0 g/L de fitoesteres y redujo a un 3,2% la cantidad de AD acumulado como subproducto (Liu et al., 2023a). Esta tiolasa FadA5 también está delecionada en el mutante de *M. neoaurum* WIII Δ CSND Δ fadA5-IV, que además sobreexpresa la aldolasa Atf1 y la hidratasa ChsH4, y llega a producir 20,7 g/L de 4-HBC a partir de 30,0 g/L de fitoesteres, disminuyendo un 56,4% su acumulación de AD (Zhao et al., 2023).

A pesar de todas estas modificaciones realizadas en el genoma, todas las micobacterias utilizadas para producir 4-HBC siguen acumulando algo de AD como subproducto. La estrategia elegida para evitarlo en *M. neoaurum* y *M. fortuitum* fue la mutación del gen *fadA5*, consiguiendo una pequeña disminución en la acumulación de AD, pero no eliminándolo por completo (Liu et al., 2023a; Zhao et al., 2023). Este resultado sugiere la existencia de alguna isoenzima capaz de suplir la actividad de la enzima eliminada y pone de manifiesto la gran redundancia génica que hay en el metabolismo del colesterol en estos microorganismos. Como alternativa en *M. smegmatis*, la deleción de la deshidrogenasa ChsE2-ChsE1 (MSMEG_5993-5994), involucrada en la siguiente reacción tras la tiolasa FadA5 (Fig. I4), podría bloquear la formación de AD. Esto se podría combinar además con la sobreexpresión de la aldolasa msSal y la reductasa msRed junto con la hidratasa msShy, ya que Zhao et al. (2023) han comprobado en *M. neoaurum* que su proteína homóloga ChsH4 es esencial en el metabolismo del 4-HBC y que su sobreexpresión en la cepa WIII Δ CSND Δ fadA5-IV disminuye la producción de AD como subproducto.

1.2. Ampliación de la herramienta genética *Golden Standard* para micobacterias y su aplicación en la producción de progesterona

La modificación genética de micobacterias, y específicamente de *M. smegmatis*, se lleva realizando desde la década de los 90 (Snapper et al., 1990), sin embargo, las herramientas moleculares utilizadas en estos microorganismos siguen estando muy limitadas al uso de unos pocos plásmidos y sistemas de clonaje tradicionales (Chang et al., 2009; Snapper et al., 1990; Stover et al., 1991).

Para mejorar esta situación, Arnold et al. (2018) desarrollaron una colección de vectores con utilidad en micobacterias para delección génica, complementación y producción de proteínas. Este sistema se basa en el uso de la enzima de restricción de tipo IIS *SapI*, permitiendo el clonaje de un fragmento en un denominado ‘vector entrada’ y su posterior transferencia al vector de interés. Esta colección de vectores además incluye una variedad de sistemas con promotores de distinta fuerza de expresión, constitutivos o inducibles, con sistemas integrativos o replicativos y con distintos marcadores de selección. Oliveira et al. (2019) por otro lado aplicaron la tecnología *BioBrick* a *M. bovis*, creando el plásmido pUP500 replicativo en *E. coli* y *M. bovis* y construyeron una colección de partes como promotores, genes reporteros, marcadores de selección y secuencias reguladoras. En este caso los clonajes se basan en enzimas de restricción de tipo IIP como *EcoRI* o *XbaI*.

Sin embargo, todos estos sistemas presentan aún varios problemas. El primero de ellos es la falta de intercambiabilidad entre las distintas colecciones. Debido a la diferencia en la forma de clonaje, las partes de una colección no se pueden usar en la otra, obligando a adaptar los fragmentos a cada una de las herramientas. Además, los vectores finales no presentan un amplio rango de hospedador, ya que se centran únicamente en *E. coli* y las micobacterias, y esto delimita mucho la aplicabilidad de las herramientas moleculares. El sistema *Golden Standard* desarrollado por Blázquez et al. (2023) soluciona estos inconvenientes al estandarizar el sistema de clonaje con las enzimas de restricción IIS *BsaI* y *BbsI* utilizadas en otros sistemas como JUMP o *Cyano Gate* y además amplía su aplicación a un gran número de bacterias al incorporar la modularidad de los pSEVA, permitiendo cambiar el origen de replicación. En la colección actual solo se encuentran adaptados al GS los vectores con los orígenes de replicación RK2, pBBR1 y pUC, cuya aplicación se ha comprobado en α -, β -, γ - y δ -proteobacterias (Blázquez et al., 2023).

En esta Tesis Doctoral se ha conseguido expandir el uso del GS como herramienta molecular para las micobacterias gracias al *ori* RSF1010. Tras la ‘domesticación’ del origen de replicación se han construido diferentes vectores que permiten la construcción de unidades transcripcionales simples o circuitos genéticos que replican y son funcionales en *M. smegmatis*. Además, se ha ensayado la colección existente de promotores, RBS, terminadores y sistemas de inducción e incluso se ha aumentado con partes funcionales para micobacterias.

Todos los promotores de la colección de Anderson o de Zobel probados en *M. smegmatis* fueron funcionales y la mayor señal se observó con los promotores J23100, J23102, BG17 y BG37. Estos niveles de expresión coinciden con los observados en *E. coli* con el *ori* RSF1010 y con el *ori* RK2 en *E. coli* y *Pseudomonas putida* (Blázquez et al., 2023). Todos ellos tienen en común que las secuencias de las cajas -10 y -35 se encuentran muy conservadas respecto a la secuencia consenso de *E. coli*, facilitando la transcripción por la ARN polimerasa (Tabla D1). Para ampliar esta colección de partes, se decidió probar los promotores Phsp60, PhspX y Pimyc, todos ellos ensayados previamente en otras micobacterias (Chang et al., 2009; Oliveira et al., 2019; Stover et

al., 1991). El promotor Phsp60 se encuentra en el plásmido pMV261 y es comúnmente utilizado en *M. smegmatis*, *M. bovis* o *M. neoaurum* para la expresión heteróloga de proteínas (Li et al., 2024; Liu et al., 2022; Reyes-González et al., 2025), siendo considerado un promotor fuerte en *M. bovis* (Bastos et al., 2009; Oliveira et al., 2019). Sin embargo, como se observa en la Figura R1.5 y como señalan Kanno et al. (2016) y Oliveira et al. (2019), este promotor resultó ser bastante débil tanto en *M. smegmatis* como en *M. bovis*. El promotor Pimyc presentó una actividad menor que Phsp60, igual que se observó Chang et al. (2009) que lo clasifican como el promotor con menor expresión de los 3 ensayados. Por otro lado, la falta de actividad de PhspX puede deberse a que este promotor se ha descrito como inducible en *M. bovis* BCG cuando infecta macrófagos o en condiciones de bajo oxígeno (Dokladda et al., 2015; Oliveira et al., 2019), que son situaciones no testadas en nuestro ensayo.

Tabla D1. Secuencias de los promotores ensayados en *M. smegmatis*. La secuencia consenso de *E. coli* de la caja -35 se encuentra marcada en naranja, mientras que la caja -10 está marcada en verde. Los nucleótidos de las cajas -35 o -10 que no coinciden con la secuencia consenso están marcados en azul.

Promotor	Secuencia
Consenso	TTGACANNNNNNNNNNNNNNNNNNTATAATNNNNNN
J23100	TTGACGGCTAGCTCAGTCCTAGGTACAGTGCTAGC
J23102	TTGACAGCTAGCTCAGTCCTAGGTACTGTGCTAGC
Pem7	TTGACAATTAATCATCGGCATAGTATATCGGCATA
BG17	TTGACAACACTCGAAAAGCCGAGTATAATCAGATG
BG37	TTGACATGTCAATTTTTATGTTGTATAATATAACT
Phsp60	TTGCACTCGGCATAGGCGAGTGCTAAGATAACGT
PhspX	TTCTGTCCCTAGCCCTGGCCACGATGGGCTGGTCG
Pimyc	TTTCGCAGTCGAGGGCTCTCGTGTAGCCTGGGCGA

Los niveles de la traducción generados por los RBS bicistrónicos en *E. coli* y *M. smegmatis* fueron superiores a los obtenidos con los RBS simples, como se ha observado en *P. putida* (Blázquez et al., 2023; Iverson et al., 2016; Mutalik et al., 2013). Este aumento de la traducción gracias a los RBS bicistrónicos se atribuye a la actividad helicasa de los ribosomas, que empiezan traduciendo el péptido líder desde el primer RBS y así eliminan las estructuras secundarias del ARNm que puedan bloquear el acceso al segundo RBS encargado de la traducción del gen de interés (Claassens et al., 2019; Takyar et al., 2005).

Los terminadores ensayados muestran una gran diferencia entre bacterias. Con la excepción del terminador T500, el resto mostraron una gran efectividad en *E. coli*, mientras que en *M. smegmatis* solo el terminador B1006 y el terminador de *rnpB*

presentaron una efectividad mayor del 50%. Estos datos en conjunto muestran las grandes diferencias que hay en la actuación de distintos elementos genéticos en función del microorganismo en el que se encuentran y, además, no solo recalcan la importancia de caracterizar cada parte en cada ambiente celular, sino también la necesidad de construir una librería amplia en la que haya elementos disponibles y útiles para cualquier especie.

Finalmente, la herramienta GS cuenta con distintos sistemas inducibles testados en *E. coli* y *P. putida* (Blázquez et al., 2023). Estos sistemas son los pares LacI/P_{trc}, RhaSR/PrhaBAD y XylS/P_m y son esenciales para una expresión controlada en el tiempo del gen de interés, ya que permite trabajar con genes que pueden resultar tóxicos o deletéreos para la célula. Estos sistemas aplicados en un vector con el *ori* RSF1010 mostraron el mismo comportamiento en *E. coli* que cuando se ensayaron con el *ori* RK2 por Blázquez et al. (2023). Sin embargo, ninguno de ellos fue funcional en *M. smegmatis*, ya sea por falta de represión o por falta de expresión del gen reportero. Considerando que estos sistemas se obtuvieron de bacterias Gram negativas como *E. coli* y *P. putida* (Egan y Schleif, 1993; Gawin et al., 2017; Kelly et al., 2016), cabe esperar que no presenten aplicabilidad en bacterias Gram positivas, especialmente por los distintos factores σ de cada microorganismo (Souza et al., 2014). Para conseguir un sistema inducible en micobacterias, se replicó el sistema TetR/P_{tet} empleado por Arnold et al. (2018) para *M. smegmatis*. A pesar de que los elementos TetR/P_{tet} se obtuvieron en *E. coli* (Bertram y Hillen, 2008), para asegurar la funcionalidad de este sistema en micobacterias se expresó *tetR* bajo el promotor débil P_{imyc}, previamente ensayado en *M. smegmatis*, y se optimizaron los codones del gen *tetR* para facilitar su traducción. De esta manera, el nuevo sistema resultó ser inducible en presencia de aTC, constituyendo el primer sistema inducible para micobacterias de la colección GS.

El desarrollo de la nueva colección de vectores con el *ori* RSF1010 para GS facilita su aplicación en un mayor número de especies, incluidas las micobacterias. Esta nueva herramienta molecular permitirá la construcción de distintas unidades transcripcionales y circuitos genéticos con aplicación en distintos procesos, como por ejemplo la producción de progesterona en *M. smegmatis*, uno de los principales objetivos de esta Tesis Doctoral.

Al inicio de esta Tesis Doctoral, solo se había descrito la producción microbiana de pregnenolona o progesterona en bacterias como *E. coli* (Efimova et al., 2019; Makeeva et al., 2013), *B. megaterium* (Gerber et al., 2015) y *M. neoaurum* (Liu et al., 2022) y en las levaduras *Saccharomyces cerevisiae* (Duport et al., 1998) y *Yarrowia lipolytica* (Zhang et al., 2019). Todas estas producciones se fundamentan en la expresión del CYP450scs bovino, denominado CYP11A1, junto con su adrenodoxina Adx y la reductasa AdxR, formando el complejo CH/L. La producción de estos compuestos en *E. coli* presenta una ventaja, ya que esta bacteria carece de CYP450 propios, favoreciendo el estudio de los CYP450 expresados heterológamente. Sin embargo, al inducir la expresión del operón del sistema CH/L con IPTG, se produjeron unos niveles de progestágenos que variaron entre los 0,42 y 0,62 mg/L a las 24 h a partir de 193 mg/L de colesterol (Efimova et al., 2019; Makeeva et al., 2013). Esta producción es mucho mayor en *B. megaterium*,

llegando a los 95 mg/L de pregnenolona desde 115 mg/L de colesterol a las 48 h gracias a la presencia de gránulos del bioplástico poli(3-hidroxibutirato) (PHB), que actúan como reserva del sustrato y unen las proteínas heterólogas a su membrana (Gerber et al., 2015). En micobacterias se consiguió una mayor concentración de progesterona en un mutante de *M. neoaurum* productor de 4-HBC expresando en formato operón el sistema CH/L de mamíferos unido a nanopartículas de indio (In), que actúan como donador de electrones ante una fuente de luz. Esta cepa produjo hasta 235 mg/L a partir de 5 g/L de colesterol tras 5 días, aunque también acumuló unos 200 mg/L de 4-HBC (Liu et al., 2022). La aproximación utilizada en levaduras consiste en una producción endógena del sustrato (ergosta-5-enol y campesterol) gracias al anabolismo del ergosterol y su posterior transformación en pregnenolona por el sistema CH/L. Se obtuvieron así 60 mg/L y 78 mg/L de este esteroide en *S. cerevisiae* y *Y. lipolytica*, respectivamente, aunque también se detectó la presencia de ésteres de pregnenolona en el medio (Duport et al., 1998; Zhang et al., 2019).

Todos estos datos apuntan a la necesidad de optimizar la expresión heteróloga del sistema CH/L para mejorar la producción de pregnenolona o progesterona en microorganismos. En esta Tesis Doctoral se decidió abordar este objetivo en *M. smegmatis*, ya que es una cepa excelente para la biotransformación de esteroides en esteroides con interés industrial, y además no se había probado todavía para este propósito al inicio de la misma. Además, se decidió utilizar en esta cepa el CYP450scc bacteriano descrito en la patente EP2386634A1 (Yamagishi et al., 2011). Este CYP450scc proviene de la α -proteobacteria *N. aromaticivorans* y es capaz de cortar la cadena lateral del colesterol y otros esteroides para producir pregnenolona, de igual forma que cataliza esta reacción el sistema CH/L de mamíferos. Se trata del primer CYP450scc bacteriano descrito hasta la fecha, por lo que supone de gran interés conocer su actividad *in vivo*. El CYP11A1 ha tenido que ser modificado varias veces para mejorar su expresión y solubilidad en bacterias, eliminando parte del N-terminal y el giro F-G para liberar el CYP11A1 de la membrana (Janocha et al., 2011). Además, la expresión de la AdxR en bacterias es también un desafío (Gerber et al., 2015). La existencia de un CYP450scc de origen bacteriano, con sus respectivas Fdx y FdxR, debería suponer por tanto una ventaja a la hora de expresar y utilizar estas enzimas en otras bacterias.

La sobreexpresión del sistema CH/L bacteriano en la cepa mejorada Δ Cyp125A3- Δ 142A2- Δ 125A4 de *M. smegmatis* permitió la producción de progesterona a partir de fitoesteroides y confirmó que, igual que se ha observado en el sistema CH/L de mamíferos (Liu et al., 2022), es necesaria la expresión conjunta del CYP450scc, la Fdx y la FdxR para conseguir actividad. La cantidad de progesterona obtenida fue de 5,1 mg/L a partir de 500 mg/L de fitoesteroides a las 48 h, unos niveles bastante por debajo de los obtenidos en otras bacterias como *M. neoaurum* o incluso *B. megaterium*. Estos resultados se pueden comparar con los datos recientes publicados por Karpov et al. (2024) en *M. smegmatis*. Karpov et al. (2024) expresaron el sistema CH/L de mamíferos en la cepa silvestre de *M. smegmatis* en formato operón y bajo un sistema inducible por acetamida y llegaron a producir 19,5 mg/L de progesterona desde 1 g/L de fitoesteroides. Esta diferencia puede deberse a la actividad propia de cada sistema CH/L, de forma que el

mamífero sería más eficiente respecto al bacteriano, o también puede ser debido a los vectores replicativos y las unidades transcripcionales construidas para ambas cepas. En nuestro caso los genes se encuentran expresados bajo el promotor Phsp60, que tiene una actividad débil, y por tanto los niveles de expresión podrían ser menores que los alcanzados cuando se usa el sistema inducible por acetamida.

Los datos representados en la Figura R1.12 muestran que una versión deshidrogenada de la progesterona, la 1,2-dehidroprogesterona, también se acumula en el tiempo, de igual forma que se observa este intermediario en el catabolismo de la progesterona de *C. tarдаugens* estudiado en el Capítulo 2. Este compuesto es la primera vez que se observa su producción en *M. smegmatis*, ya que Karpov et al. (2024) no indican la presencia de este esteroide en sus ensayos. Este compuesto se debe producir por la acción de alguna de las 6 KstD presentes en el genoma de la micobacteria (Galán et al., 2017), de forma que la delección o la sobreexpresión de estas deshidrogenasas disminuirían o aumentarían la cantidad de 1,2-dehidroprogesterona obtenida. La 1,2-dehidroprogesterona puede considerarse también un esteroide con interés farmacéutico, ya que no solo sirve como punto de partida para sintetizar distintos progestágenos, sino que también se ha comprobado su actividad antiviral en células MDCK frente a la cepa A/WSN/33 (H1N1) del virus de la gripe y presenta un IC₅₀ similar al del antiviral oseltamivir (Cheng et al., 2016; De et al., 1991).

La acumulación de progesterona también se observó en la cepa control *M. smegmatis* Δ Cyp125A3- Δ 142A2- Δ 125A4 (pSEVA65g) únicamente cuando se usó colesterol como materia prima, siendo la primera vez que se describe la producción endógena de progesterona en *M. smegmatis* y, hasta donde alcanza nuestro conocimiento, en otras micobacterias. A la hora de buscar algún CYP450 similar al CYP450scc de *N. aromaticivorans* en *M. smegmatis*, los alineamientos no superaron el 36% de cobertura ni de identidad. Esto indica que la progesterona producida puede deberse a la acción de un CYP450scc desconocido, a la actividad secundaria de otro CYP450 o por algún tipo de descarboxilación desde el 3-OPA.

A pesar de que el rendimiento de producción de progesterona en *M. smegmatis* (19 mg/g colesterol) es muy parecido al obtenido por Karpov et al. (2024) (23 mg/g), la productividad fue mucho menor (0,19 mg/L h⁻¹ frente a 0,81 mg/L h⁻¹). Para mejorar su producción, en vez de delecionar los CYP450 involucrados en el catabolismo del colesterol y evitar su degradación, Karpov et al. (2024) bloquearon el hidroxilo unido en C3 de los esteroides con un metoximetilo, de forma que el hidroxilo de C3 ya no se pudiera oxidar y por tanto la molécula no solo era peor sustrato para el metabolismo del esteroide en la bacteria, sino también para evitar su transformación en colesteno, cuya afinidad por el sistema CH/L de mamíferos es mucho menor. De esta forma, consiguieron concentraciones de hasta 461 mg/L de MOM-pregnenolona a partir de 1 g/L de MOM-colesterol. Sin embargo, esto requiere de un posterior proceso químico para recuperar la pregnenolona, lo que no cumple con el objetivo de una completa producción biológica de este compuesto. En *M. neoaurum* o *Y. lipolytica* se realizó un cribado previo con distintos promotores o combinaciones de Adx o AdxR para reconocer la construcción ideal para la

producción (Liu et al., 2022; Zhang et al., 2019). La expresión conjunta del sistema CH/L con la reductasa ARH1 de levaduras en *M. neoaurum* consiguió la mayor acumulación de progesterona, mientras que en *Y. lipolytica* la combinación del CYP11A1 de *Sus scrofa* junto con la Adx y AdxR de *Bos Taurus* resultó ser el mejor sistema CH/L. Además, Zhang et al. (2019) comprobaron que una mayor expresión del CYP11A1 aumentaba la concentración final de pregnenolona. Esto se debe a que la ratio de reducción del CYP11A1 es mucho más baja que la de la Adx o la AdxR ($1,86 \text{ s}^{-1}$ frente a $5,4 \text{ s}^{-1}$ y $32,2 \text{ s}^{-1}$), siendo por tanto el paso limitante en esta catálisis (Schiffler et al., 2004).

El balanceo de los niveles de expresión de estos genes constituye por tanto una aproximación necesaria para mejorar la producción de progesterona. El cribado de distintos promotores se intentó ensayar en *M. smegmatis* gracias a la ampliación del GS en micobacterias, que permite aumentar la combinatoria de promotores. Sin embargo, los transformantes obtenidos presentaban plásmidos aberrantes probablemente porque unos altos niveles de expresión del sistema CH/L resultan ser una carga metabólica excesiva para la célula.

Tanto los resultados recogidos en este apartado como las herramientas genéticas desarrolladas para modificar *M. smegmatis* están en línea con resultados previos en los que se ha propuesto como una excelente factoría celular para abordar la producción de diversos esteroides. El estudio más detallado del metabolismo de los esteroides ha permitido obtener nuevas cepas con una aplicación industrial muy prometedora al conseguir aumentar la producción de sintonas de interés, aunque en algunos casos se requiere todavía de algunas mejoras para evitar, sobre todo, la acumulación de subproductos. La progesterona es un esteroide clave cuya producción se puede conseguir a partir del intermediario 4-HBC o directamente aplicando la biotecnología de CYP450scc. Su producción microbiana a escala industrial puede ser una realidad usando como materia prima fitoesteroides obtenidos de residuos de la industria papelera o de la soja.

2. Metabolismo de la progesterona en *C. taradagens*

C. taradagens se aisló por su capacidad de degradar estrógenos y sin embargo es capaz de metabolizar además un gran número de esteroides y otros EDCs como testosterona y sus ésteres, AD, ADD, ácido cólico, progesterona, pregnenolona, colesterol y colesteno. Por este motivo, la hemos empleado en diversos estudios como cepa Gram negativa modelo para la degradación de estas moléculas (Ibero, 2022; Ibero et al., 2020, 2019). Ibero et al. (2020, 2019) describieron los *clusters* génicos de degradación de la testosterona y del estradiol en *C. taradagens*, denominados *clusters* edc y SD, respectivamente, y completaron las rutas metabólicas con ensayos enzimáticos y análisis de intermediarios. El metabolismo de la progesterona y los genes involucrados, en cambio, todavía eran desconocidos tanto en *C. taradagens* como en otras bacterias.

En el Capítulo 2 de esta Tesis Doctoral se ha estudiado el transcriptoma de *C. taradagens* en presencia de progesterona para conocer cuáles son los genes encargados

de degradar este esteroide, especialmente aquellos involucrados en la degradación de su cadena lateral. Este análisis reveló la existencia del *cluster* pdc (por sus siglas, *progesterone degradation cluster*) en el que se encuentran las monooxigenasas BVMO (EGO55_13845) y LLM (EGO55_13860) y los reguladores EGO55_13850 y EGO55_13855. Este *cluster* también se encuentra en otras bacterias degradadoras de esteroides como *N. aromaticivorans* o *C. estronivorus*, en la que además también se conserva el *cluster* edc y SD descritos por Ibero et al. (2020, 2019). Estos datos, junto con el hecho de que los *clusters* pdc, edc y SD se encuentran contiguos en una región de 75,5 kb, sugiere una gran conservación de la organización génica a lo largo de la evolución.

Las monooxigenasas BVMO son una familia de proteínas bien caracterizadas que catalizan una reacción de Baeyer-Villiger sobre un gran número de sustratos con un grupo ceto, como la ciclobutanona y la fenilacetona, usando FAD como cofactor (Bisagni et al., 2014; Franceschini et al., 2012; Miyamoto et al., 1995). Como se ha comprobado en este trabajo, la BVMO EGO55_13845 transforma la progesterona en acetato de testosterona consumiendo poder reductor, de igual forma que se ha observado en la BVMO de *R. rhodochrous* y cuya estructura 3D se encuentra resuelta (Franceschini et al., 2012). Otras BVMOs descritas en hongos catalizan también la D-lactonización de esteroides como progesterona, testosterona y dehidroepiandrosterona, convirtiéndolos en testololactona (Ferroni et al., 2016; Javid et al., 2018). Al mismo tiempo, la monooxigenasa LLM EGO55_13860 cataliza la misma reacción que BVMO y transforma la progesterona en acetato de testosterona, pero en este caso la enzima requiere de un componente reductor externo y de FMN para la reacción. Por esta razón, las enzimas LLM se encuentran menos estudiadas. Los únicos ejemplos descritos son el complejo AbmE2/AbmZ de *Streptomyces koyagensis* SCSIO 5802, encargado de inactivar el antimicrobiano abisomicina 2 al transformarlo en neoabisomicina B, y las enzimas 2,5-dicetocanfano-1,2-monooxigenasa y 3,6-dicetocanfano-1,6-monooxigenasa de *P. putida* ATCC 17453 (Isupov et al., 2018; Iwaki et al., 2013). De esta forma, esta LLM EGO55_13860 es la siguiente enzima de esta familia en ser estudiada, aunque todavía se desconoce cuál es el componente reductor que actúa junto a ella.

En este *cluster* pdc se detectó la presencia de dos reguladores, EGO55_13850 y EGO55_13855 pertenecientes a la familia de represores TetR/AcrR y a la familia de reguladores LuxR, respectivamente. En *C. testosteroni* TA441 se ha descrito el regulador transcripcional TesR, de la familia de LuxR, que participa en la regulación de la degradación de la testosterona induciendo distintos genes (Horinouchi et al., 2004b), lo que contrasta con la posible función represora de EGO55_13850. Sería interesante por tanto estudiar la actividad de estos reguladores y su papel en la degradación de la progesterona y describir así su función inductora o represora, las secuencias de ADN operadoras y los inductores frente a los que responden.

El análisis funcional mediante mutagénesis y complementación demostró que las dos monooxigenasas son redundantes porque tienen una actividad similar, de forma que la delección de ambas enzimas generó un mutante incapaz de crecer en progesterona como

única fuente de carbono y de energía. En este mutante, además, se detectó el intermediario 1,2-dehidroprogesterona. La presencia de esta molécula también fue observada por Liu et al. (2013) al estudiar los intermediarios acumulados en la degradación aerobia de la progesterona y del norgestrel por bacterias provenientes de fangos activos, y su presencia se debe a la actividad de las deshidrogenasas KstD. Estas enzimas están muy estudiadas y se usan para obtener esteroides C19 y C21 deshidrogenados con aplicaciones farmacéuticas (García et al., 2012; Zhao et al., 2021). En *C. tardaugens* se encuentra la KstD EGO55_13510 involucrada en la degradación de la testosterona (Ibero et al., 2019), pero en su genoma se encuentran otras KstD potenciales como EGO55_01175 y EGO55_03150 que podrían catalizar esta reacción. En los análisis transcriptómicos se observó que el rango de FC de estos genes estaba entre -1 a +1, lo que sugiere que estos genes se expresan de forma similar en presencia de progesterona y de testosterona, probablemente debido a que el catabolismo de la progesterona confluye con el de la testosterona.

La redundancia enzimática de las enzimas BVMO y LLM se puede explicar por su distinta preferencia de sustratos. Esta redundancia enzimática es un fenómeno que existe en un gran número de microorganismos y por tanto debe presentar una ventaja para ellos. En el caso de la actividad de BVMO y LLM en *C. tardaugens*, esta característica aumentaría la versatilidad del metabolismo de los esteroides, ya que permite degradar otros progestágenos distintos a la progesterona y combinar su preferencia por unos u otros. En este sentido, sobre la existencia de redundancia en el metabolismo, a modo de ejemplos, en *P. aeruginosa* se ha observado que el ciclo del glioxilato está ligado al ciclo del 2-metilcitrato por la isocitrato liasa AceA, que presenta actividad secundaria con el 2-metilisocitrato y por tanto reduce la acumulación de intermediarios tóxicos cuando se muta la 2-metilisocitrato liasa PrpB (Wijaya et al., 2025). Mitchell et al. (2024) comprobaron que las proteínas de unión a penicilina (PBP) tienen redundancia funcional, pero hay una especialización entre estas proteínas según distintos estresores medioambientales.

C. tardaugens, y seguramente otras bacterias que también degradan testosterona, son capaces por tanto de metabolizar la progesterona transformándola en acetato de testosterona, cuya degradación se canaliza por el metabolismo de la testosterona. Esta relación entre ambos metabolismos confirma que el AD sería el esteroide central en el que converge el catabolismo de la mayoría de los esteroides, los ácidos biliares y los esteroides, como es el caso también del colesterol en micobacterias (Fernández-Cabezón et al., 2018b; Galán et al., 2017; Wei et al., 2014). Cuando se abren los anillos A y B del intermediario AD para entrar en el metabolismo central, éste se transforma en HIP, que es la molécula clave en la que confluye el catabolismo de los estrógenos y, por tanto, el intermediario en el que convergen todas las rutas de degradación aeróbicas de esteroides (Fig. D1).

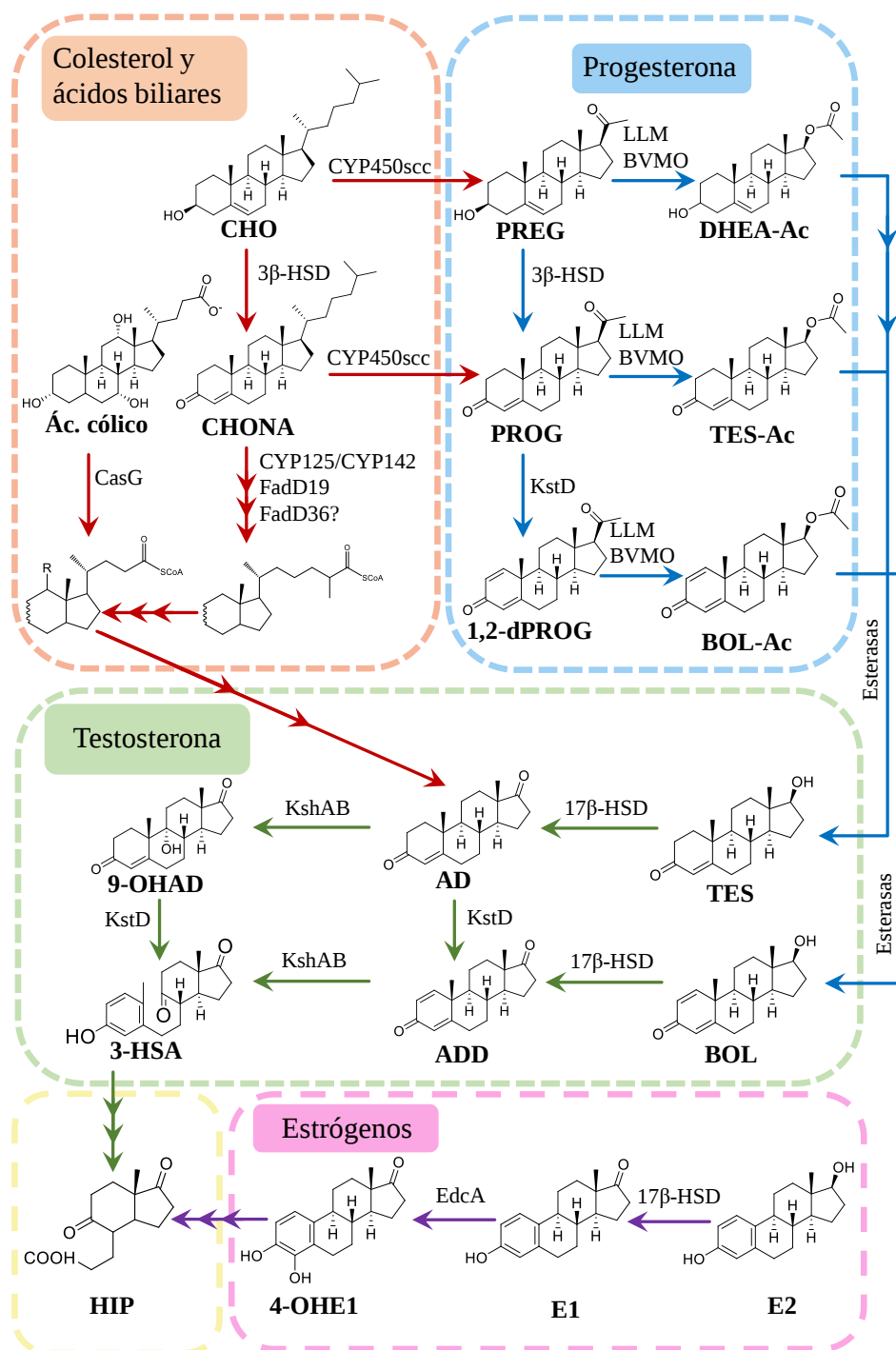


Figura D1. Esquema integrativo de las rutas de degradación aerobias de esteroides. CHO: colesterol; PREG: pregnenolona; DHEA-Ac: acetato de dehidroepiandrosterona; CHONA: colestonona; PROG: progesterona; TES-Ac: acetato de testosterona; 1,2-dPROG: 1,2-dehidroprogesterona; BOL-Ac: acetato de boldenona; 9-OHAD: 9α-hidroxi-1,4-androstadien-3,17-diona; AD: 4-androsten-3,17-diona; TES: testosterona; 3-HSA: 3-hidroxi-9,10-secoandrost-1,3,5(10)-trien-9,17-diona; ADD: 1,4-androstadien-3,17-diona; BOL: boldenona; HIP: dioxo-1,2,3,4,10,19-hexanorandrostan-5-oico; 4-OHE1: 4-hidroxiestrona; E1: estrona; E2: estradiol; CYP450scc: citocromo P450 *side chain cleavage*; LLM: luciferasa-like monooxigenasa; BVMO: Baeyer-Villiger monooxigenasa; KstD: 3-cetoesteroide-Δ¹-deshidrogenasa; KshAB: cetoesteroide-9α-hidroxilasa; 17β-HSD: 17β-hidroxiesteroide deshidrogenasa; EdcA: citocromo P450 estrona 4-hidroxilasa.

3. Metabolismo del colesterol en *C. tardaugens*

C. tardaugens NBRC 16725 es capaz de crecer en sales biliares como el colato y el desoxicolato, caracterizados por su cadena lateral ácida, y en esteroides como el colesterol y la colestenoína, que presentan una cadena lateral alifática. El metabolismo de las sales biliares en proteobacterias se encuentra muy estudiado, y sigue una ruta similar a la degradación del colesterol en la que primero se elimina la cadena lateral y, a continuación, se produce la apertura de los anillos esteroideos (Barrientos et al., 2015; Gong et al., 2012; Holert et al., 2016; Horinouchi et al., 2012). Sin embargo, el catabolismo del colesterol en Gram negativas supone todavía un campo por estudiar. Actualmente solo hay descritas dos bacterias Gram negativas que pueden mineralizar el colesterol: la cepa no clasificada 72Chol (DSM 12783) y *S. denitrificans* DSMZ 13999 (Harder y Probian, 1997; Tarlera y Denner, 2003).

La falta de genes homólogos a los involucrados en el metabolismo aeróbico del colesterol de actinobacterias o en la ruta 2,3-*seco* anaeróbica de *S. denitrificans* nos hizo plantear la posible existencia de una ruta alternativa para la degradación de este esteroide en *C. tardaugens*. Los CYP450_{scc} de *N. aromaticivorans* y *N. subterraneum* ensayados *in vitro* tienen actividad CH/L y transforman el colesterol y la colestenoína en pregnenolona y progesterona, respectivamente (Yamagishi et al., 2011) y, al mismo tiempo, Ibero (2022) describió en *C. tardaugens* el CYP2113, cuya actividad no se conoce, pero su mutación impedía el crecimiento de esta bacteria en colesterol, aunque no en colestenoína. El ensayo de este CYP2113 *in vitro* con el par Fdx/FdxR de espinaca no mostró ningún resultado de actividad en colesterol a pesar de que este par Fdx/FdxR de espinaca ha funcionado con otros CYP450 de *C. tardaugens* (Ibero et al., 2020). Por tanto, se decidió cambiar los componentes redox a unos de origen procariota. Zhang et al. (2019) realizaron una combinatoria de distintos CYP11A1 y Adx de mamíferos para mejorar la actividad del CYP450_{scc}, ya que la interacción entre CYP450_{scc} y Fdx/FdxR es esencial para la reacción. En ocasiones se suele reemplazar la actividad redox del par endógeno Fdx/FdxR por la putidaredoxina y la putidaredoxina reductasa (par CamAB) obtenidas de *P. putida* (Koga et al., 1989), pero en cambio se optó por utilizar el par Fdx/FdxR de *N. aromaticivorans* debido a que esta especie es más cercana filogenéticamente a *C. tardaugens* que *P. putida* y porque la actividad de este par Fdx/FdxR se había ensayado previamente *in vivo* en *M. smegmatis* (ver Capítulo 1). Esta nueva construcción que expresa el CYP2113 con la Fdx y la FdxR de *N. aromaticivorans* se ensayó *in vivo* en *E. coli* como está descrito en Efimova et al. (2019) y en la patente EP2386634A1 (Yamagishi et al., 2011) y no se detectó ningún nuevo compuesto en estas condiciones, ya sea la pregnenolona o algún intermediario del colesterol hidroxilado como el 22-OHCHO o el 20,22-dOHCHO, observados con el CYP11A1 (Strushkevich et al., 2011). Es necesario por tanto encontrar cuál es la reacción que cataliza el CYP2113 y cuáles son los productos que se obtienen.

Se estudiaron los perfiles de expresión de la transcriptómica realizada en presencia de colesterol de las Fdx anotadas en el genoma de *C. tardaugens*, para conocer cuáles podrían ser las candidatas para actuar con el CYP2113. La Adx de mamíferos es capaz

de actuar con distintos CYP450 relacionados con la esteroidogénesis y con el metabolismo de la vitamina D y ácidos biliares (Ewen et al., 2012). Por ello, alguna o varias de las 8 Fdx codificadas en el genoma deberían ser capaces de asistir a los CYP450 y, entre ellos, al CYP2113. De las 8 Fdx, solo *EGO55_05120* y *EGO55_19875* se encuentran sin un CYP450 cerca y, al mismo tiempo, están inducidas en presencia de colesterol, de forma que podrían estar acompañando a los 14 CYP450 inducidos en las mismas condiciones y que no presentan una Fdx al lado en el genoma. El CYP2113 tiene un FC=34 y unos TPMs de 19740, considerablemente mayores a los observados en las distintas ferredoxinas que rondan entre 40 y 6595. Sin embargo, independientemente de sus niveles de FC, podría considerarse que estas Fdx tienen una expresión basal suficiente para actuar como compañeras redox de los CYP450, gracias a que la ratio de reducción de las Fdx es casi 3 veces mayor que la de los CYP450 como el CYP11A1 (Schiffler et al., 2004). Esta misma circunstancia se observa en el caso de *EGO55_09415*, propuesta en esta Tesis Doctoral como la FdxR de *C. tardaugens*. De forma general, en las bacterias se suele encontrar una única FdxR encargada de utilizar el poder reductor del NAD(P)H para reducir las Fdx, que a su vez cede los electrones a los CYP450 u otras monooxigenasas (Giró et al., 2006; Moreno et al., 2024; Yang et al., 2010). Ibero (2022) propone 4 genes como posibles FdxR en *C. tardaugens* (*EGO55_01285*, *EGO55_05820*, *EGO55_06710* y *EGO55_09415*), pero solo la *EGO55_09415* presenta un alto porcentaje de identidad con otras FdxR como son ArR o CamA. Aunque el gen codificante se encuentre reprimido en presencia de colesterol frente a testosterona (FC=-2), sus TPMs podrían ser suficientes para expresar enzima suficiente para proporcionar el poder reductor a la Fdx. En el sistema CYP11A1/Adx/AdxR, la reducción de la AdxR es el paso más rápido, de forma que bajos niveles de esta proteína pueden suplir la demanda de poder reductor generada por las Fdx. No obstante, cabe destacar que el gen *EGO55_01285* propuesto por Ibero (2022) presenta un FC de 2 con unos valores de TPM que rondan los 6700 en presencia de colesterol, de manera que su papel en este metabolismo debería ser también tomado en cuenta. Por lo tanto, un cribado de la actividad de distintas Fdx y FdxR de *C. tardaugens* con el CYP2113 utilizando colesterol como sustrato permitiría identificar el mejor conjunto CYP450/Fdx/FdxR implicado en el catabolismo del colesterol.

Se ha comprobado que la mutación del CYP2113 impide el crecimiento de *C. tardaugens* en colesterol, sin embargo, este mutante todavía es capaz de utilizar la colestenoína como fuente de carbono y de energía. Esto sugiere la posible existencia de otro CYP450 que actuaría sobre la colestenoína cortando su cadena lateral y produciendo progesterona. En este sentido, se sabe que el rendimiento del CYP11A1 de mamíferos se reduce al 3% cuando el sustrato de reacción es la colestenoína (Morisaki et al., 1982), por lo que no se puede descartar que bacterias como *C. tardaugens* hayan evolucionado para adquirir diferentes CYP450 que les permitan metabolizar con más eficiencia un mayor número de esteroides. De entre los 7 CYP450 más inducidos en presencia de colesterol, las predicciones de su estructura tridimensional apuntan a que *EGO55_10220*, *EGO55_14410* y *EGO55_03010* podrían unir un ligando muy apolar como el colesterol en una conformación que aproxime el C22 al grupo hemo. Esta orientación es la misma

que se ha descrito en el cristal del CYP11A1 cuando une el intermediario 22-OHCHO (Mast et al., 2011). En este CYP450scc se ha observado que el intermediario hidroxilado se une con más fuerza al centro activo que el colesterol y se ha propuesto que el desplazamiento del sustrato para hidroxilar el C20 y formar 20,22-dOHCHO se desencadena al interaccionar con la Adx. Cabe destacar además que el gen *EGO55_10220* se ubica al lado del *cluster* cdc descrito por Ibero (2022), donde se encuentra el gen *EGO55_10200* que codifica el CYP2113, por lo que sería interesante estudiar la actividad de este CYP450 y confirmar si la colesteno es su sustrato preferente. Resulta interesante que los genes de este *cluster* cdc, incluyendo el *EGO55_10220*, presentan homólogos en otros microorganismos, pero su estructura génica es única de *C. taradagens*.

De todos modos, salvo el CYP450 *EGO55_14220* cuya predicción estructural no permite acoplar el colesterol en su centro activo, los CYP450 *EGO55_03005* y *EGO55_04870* tampoco se pueden descartar como potenciales CYP450scc, ya que en este caso sus estructuras tridimensionales sí permiten unir colesterol, aunque sea de forma distinta a la observada en los cristales de CYP11A1.

El gen *EGO55_13890* que codifica una oxidorreductasa es uno de los más inducidos en *C. taradagens* en presencia de colesterol y se encuentra en una región ubicada cercana al *cluster* pdc con un FC entre 2 y 57. Ibero (2022) deletionó esta oxidorreductasa en *C. taradagens*, pero no observó ningún cambio fenotípico cuando el mutante se cultivó en colesterol, demostrando que no es un gen esencial en el catabolismo del colesterol en esta bacteria. Sin embargo, este gen y la región *EGO55_13865-13930* podrían estar implicados en la degradación de algún subproducto de este metabolismo. Cuando el CYP450scc rompe la cadena lateral del colesterol, libera pregnenolona y el aldehído isocaproal, cuyo catabolismo se desconoce aún. Como las enzimas implicadas en el catabolismo de la progesterona ya se conocen y no involucra a esta región de genes inducidos en colesterol, se podría suponer que esta oxidorreductasa *EGO55_13890* actúa sobre el isocaproal oxidándolo al ácido isocaproico. Este ácido se podría activar a continuación por la CoA transferasa *EGO55_13905* y entrar en una β -oxidación por su estructura similar a un ácido graso. En este proceso actuaría también la enoil-CoA hidratasa *EGO55_13910* en uno de los ciclos, posiblemente junto con las acil-CoA deshidrogenasas *EGO55_13920* y *EGO55_13925*, que se encuentran en esta región, aunque con un FC de -1 y -2, respectivamente.

Sorprendentemente, ningunos de estos genes que codifican Fdx, FdxR y CYP450 se encuentran inducidos en presencia de progesterona frente a testosterona. Esto sugiere que la progesterona, aunque es un producto intermediario del colesterol, no sería un inductor de esta ruta y solo es capaz de inducir la expresión del *cluster* pdc (Fig. D2). Por tanto, sería el colesterol o algún intermediario entre el colesterol y la progesterona (22-OHCHO o 20,22-dOHCHO) las moléculas inductoras del *cluster* cdc. Esta situación ocurre también con la región anexa al *cluster* pdc (*EGO55_13865-EGO55_13930*). Estos genes se encuentran inducidos en presencia de colesterol, pero no con progesterona, de forma que serían el colesterol o algún derivado de su metabolismo como el isocaproal los inductores de su expresión.

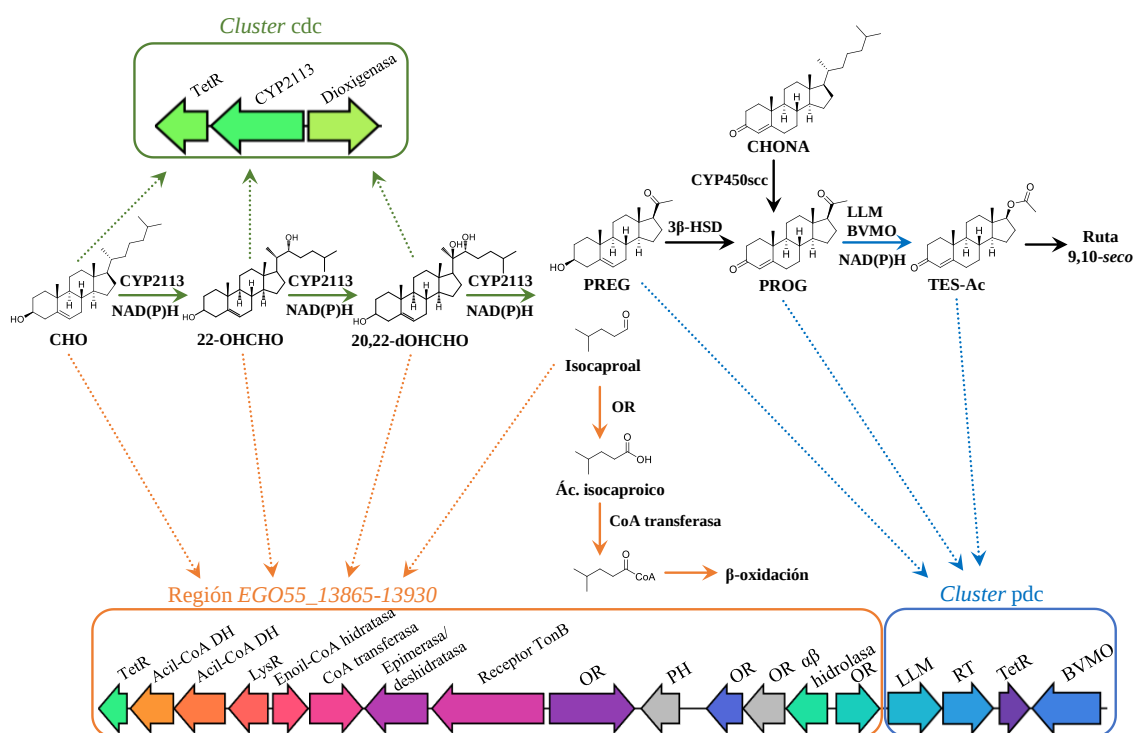


Figura D2. Esquema de la ruta propuesta de degradación del colesterol en *C. tardagens*. Los posibles efectos inductores de los compuestos sobre los genes catabólicos están marcados con flechas discontinuas. CHO: colesterol; 22-OHCHO: 22R-hidroxicolesterol; 20,22-dOHCHO: 20R,22R-dihidroxicolesterol; PREG: pregnenolona; PROG: progesterona; CHONA: progesterona; TES-Ac: acetato de testosterona; cdc: *cholesterol degradation pathway*; pdc: *progesterone degradation pathway*; DH: deshidrogenasa; OR: oxidorreductasa; PH: proteína hipotética; RT: regulador transcripcional; LLM: luciferasa-like monooxigenasa; BVMO: Baeyer-Villiger monooxigenasa; CYP450scc: citocromo P450 *side chain cleavage*.

A pesar de no conocer la actividad del CYP2113 o de algún otro CYP450scc en *C. tardagens*, la ausencia de crecimiento del doble mutante Δ LLM-BVMO en colesterol confirma la relación entre el catabolismo del colesterol y la progesterona. Esta es la primera vez que se describe *in vivo* este tipo de metabolismo en bacterias y confirma que la 1,2-dehidroprogesterona es un intermediario de la ruta, que a continuación se degradaría por la ruta 9,10-seco (Fig. D2). A diferencia del mutante simple Δ CYP2113, que no crece en colesterol, pero sí en colesteno, el doble mutante Δ LLM-BVMO no es capaz de crecer ni en colesterol ni en colesteno como única fuente de carbono y de energía hasta que no se produce una adaptación pasadas las 192 h de cultivo que le permite crecer hasta alcanzar una $DO_{600}=0,8$ (datos no mostrados). Este comportamiento adaptativo sucede tanto con colesterol como con colesteno. Estos resultados sugieren que el metabolismo de la colesteno también está ligado al catabolismo de la progesterona y utiliza la misma ruta de degradación.

Este tipo de adaptaciones para metabolizar un esteroide que *a priori* no puede ser utilizado por la bacteria ya se ha observado en *M. smegmatis* por Fernández-Cabezón et al. (2018b). Esta bacteria no es capaz de crecer en presencia de ADD como única fuente de carbono y de energía hasta su adaptación pasadas 168 h, cuando la mutación por

pérdida de función del represor PadR (*MSMEG_2868*) permite la expresión constitutiva del denominado *cluster* C-19+, que codifica genes catabólicos para la degradación de esteroides C19. Esta nueva cepa mutante adaptada a ADD a su vez es capaz de adaptarse y crecer en presencia de progesterona pasadas 168 h, mientras que la cepa WT no (datos no mostrados), indicando que esta primera adaptación genotípica permite a la cepa metabolizar nuevos compuestos.

La explicación de las observaciones anteriores surge de la comprobación de que la cepa Δ LLM-BVMO es capaz de crecer en progesterona pasadas 264 h (datos no mostrados). Estos resultados respaldan la hipótesis de que después de un periodo de adaptación podría expresarse otra enzima que sustituye a las enzimas mutadas. Esta hipótesis está respaldada porque existe codificada en el genoma otra enzima BVMO EGO55_06865 que presenta una similitud del 53% con la BVMO EGO55_13845 (ver Capítulo 2), y podría ser la responsable de permitir que el doble mutante termine creciendo en presencia de colesterol, colesteno o progesterona tras un periodo de activación de la expresión de esta enzima. El gen *EGO55_06865* presenta un FC de 2 en presencia de colesterol, mientras que en presencia de progesterona muestra un bajo nivel de expresión (FC=1 y TPMs=50 frente a FC=139 y TPMs=4330 de su homólogo *EGO55_13845*), lo que sugiere que la cepa primero debe adaptarse para producir niveles suficientes de esta enzima y metabolizar los progestágenos. La presencia de actividad de esta BVMO explicaría la acumulación transitoria de 1,2-dehidroprogesterona observada en la Figura R3.5 y, al mismo tiempo, justificaría que al cabo de 336 h la cepa doble mutante Δ LLM-BVMO alcance una DO₆₀₀ similar a la cepa WT (Fig. R3.4), ya que los metabolitos de la ruptura del colesterol se acabarían metabolizando.

La manera en la que esta adaptación se produce no se conoce, pero es posible que la adaptación se explique por una desregulación génica como la que se ha observado en *M. smegmatis* para el AD (Fernández-Cabezón et al., 2018b), ya sea porque se produce una mutación en un regulador que por activación o desrepresión aumenta la expresión del gen *EGO55_06865* o porque se produce una mutación en el mismo gen que aumenta la actividad de la enzima, o porque fenotípicamente se activa algún sistema de regulación pleiotrópico que aumenta indirectamente la expresión del gen. Habría que realizar estudios ómicos para determinar cuál de estas explicaciones es la verdadera.

La conexión entre el metabolismo del colesterol y el de la progesterona permite la utilización de *C. tardagens* Δ LLM-BVMO como factoría celular para abordar la producción de progestágenos. En los ensayos de *resting cells* con el doble mutante Δ LLM-BVMO se llegó a alcanzar 0,31 mg/L de 1,2-dehidroprogesterona, concentraciones similares a las observadas en *E. coli* expresando el sistema CH/L de mamíferos (Efimova et al., 2019; Makeeva et al., 2013), pero mucho menores que cualquier otra cepa descrita en este trabajo o en bibliografía (Duport et al., 1998; Gerber et al., 2015; Liu et al., 2022; Zhang et al., 2019). Un deficiente transporte del colesterol al interior celular puede ser una de las razones de un rendimiento tan bajo y explicaría además que tras 30 h de biotransformación todavía queden sin utilizar 232 mg/L de este sustrato en el medio, lo que contrasta con los 16 mg/L de colesterol restantes en la cepa

de *M. smegmatis* $\Delta Cyp125A3-\Delta 142A2-\Delta 125A4$ (p65_Phsp60_CYP_Fdx_ArR). En *M. smegmatis* se encuentra el operón *mce4* encargado de importar colesterol mediante transportadores ABC y otras proteínas asociadas, sin embargo, este sistema no se encuentra en *C. tardaugens* (García-Fernández et al., 2017b). La capa de lipopolisacáridos presente en la membrana externa de las bacterias Gram negativas y la ausencia de ATP en el espacio periplásmico constituye un obstáculo a la hora de transportar compuestos apolares al interior celular (Maher y Hassan, 2023; Plésiat y Nikaido, 1992). Se ha propuesto que *S. denitrificans* utiliza transportadores de la familia de FadL para superar este problema y, además, se ha comprobado que esta cepa mejora el importe de colesterol en presencia de CDX como molécula transportadora (Lin et al., 2015). Sin embargo, al no existir homólogos del sistema tipo FadL en *C. tardaugens*, Ibero (2022) propone una serie de proteínas de membrana de la familia YeeE/YedE que podrían estar implicadas en este proceso. En la región *EGO55_13865-EGO55_13930* se encuentra un receptor TonB con un FC de 14. Estos receptores parecen ser capaces de transportar moléculas apolares y por tanto no se podría descartar la implicación de este gen en el transporte (Ibero, 2022; Lin et al., 2015). Sin embargo, todavía se desconoce el mecanismo exacto de transporte de colesterol al interior celular en *C. tardaugens* e incluso se ha observado que esta bacteria no es capaz de crecer si el esterol se proporciona sin disolver o si está disuelto en detergentes como el tyloxapol, indicando que las CDX juegan un papel importante para el transporte de este esterol, igual que en *S. denitrificans*.

En cuanto a la producción de progesterona, la productividad de la cepa de *C. tardaugens* $\Delta LLM-BVMO$ es menor que la de la cepa de *M. smegmatis* $\Delta Cyp125A3-\Delta 142A2-\Delta 125A4$ (p65_Phsp60_CYP_Fdx_ArR) a las 7 h (0,044 mg/L h⁻¹ frente a 0,13 mg/L h⁻¹), y llega a ser un orden de magnitud menor a las 24 h (0,01 mg/L h⁻¹ frente a 0,19 mg/L h⁻¹). En *C. tardaugens* destaca la ausencia de progesterona en el medio, siendo la 1,2-dehidroprogesterona el único compuesto detectado. El gen *kstD* cuyo producto es el encargado de deshidrogenar el AD y otros esteroides en *M. smegmatis* se encuentra reprimido en el regulón *kstR*, aunque hay otras 5 KstD anotadas en su genoma (Galán et al., 2017). Es de esperar por tanto que no se observe actividad KstD deshidrogenasa hasta la inducción del regulón, mientras que una expresión constitutiva de este gen en *C. tardaugens* explicaría la acumulación de 1,2-dehidroprogesterona. En función del producto final requerido se podría por tanto eliminar la actividad KstD en *C. tardaugens* para obtener solo progesterona o expresar la enzima KstD constitutivamente en *M. smegmatis* para conseguir más 1,2-dehidroprogesterona.

El doble mutante $\Delta LLM-BVMO$ tiene potencial para poder ser utilizado como productor de progestágenos, pero presenta varias desventajas respecto a otras bacterias utilizadas a nivel industrial como las cepas de *Mycolicibacterium*. Centrándose en condiciones más industriales, la CDX no se usa como solvente para los esteroides debido a su alto coste, y se prefiere utilizar aceites en los que se disuelve los fitoesteroides, la materia prima más utilizada en las biotransformaciones. *C. tardaugens* no es capaz de utilizar los fitoesteroides como fuente de carbono y de energía y, dado que en los ensayos *in vivo* e *in vitro* con los CYP450scs bacterianos se ha detectado actividad con fitoesteroides como β -sitosterol, estigmasterol o campesterol, se puede pensar que el

problema radica, una vez más, en su transporte. Esto dificulta el uso de sustratos alternativos más baratos para la industria. La fuente de energía utilizada en el proceso también influye en el rendimiento final ya que el sistema CH/L requiere 3 ciclos de reducción para cortar la cadena lateral del colesterol. Igual que en los ensayos con la cepa de *M. smegmatis* se ha demostrado que la suplementación con glicerol proporciona mejores resultados en la producción de progesterona, sería necesario realizar un cribado con *C. tardaugens* para estudiar la fuente de energía óptima para estos procesos. Esta bacteria no es capaz de metabolizar la glucosa ni el glicerol, por lo que se decidió emplear un medio rico como el NB para los ensayos en *resting cells*. Sin embargo, este medio está compuesto principalmente por aminoácidos y no parecen ser el sustrato idóneo para producir 1,2-dehidroprogesterona. *C. tardaugens* muestra un crecimiento óptimo en presencia de estrógenos y andrógenos comparado con otras fuentes de carbono más comunes como el piruvato. El metabolismo de estos compuestos genera varias moléculas de acetil-, propionil- y succinil-CoA que proporcionan una gran cantidad de poder reductor a la célula (Ibero et al., 2020, 2019). Por último, como se observa en la Figura R3.6, la cepa Δ LLM-BVMO sigue siendo capaz de metabolizar parte del producto acumulado, ya sea por redundancia enzimática posiblemente debida a la BVMO EGO55_06865 o por otras modificaciones que sufra la 1,2-dehidroprogesterona, lo que disminuye el rendimiento del proceso a menos que se eliminen también estas actividades.

En resumen, los resultados obtenidos han confirmado que el colesterol se degrada en *C. tardaugens* por una ruta alternativa no descrita hasta la fecha en bacterias en la que la progesterona/pregnenolona juegan el papel de intermediarios. Este metabolismo es completamente distinto a la conocida ruta de degradación del colesterol en actinobacterias y presumiblemente comienza con la acción de un CYP450_{scc}. El papel de CYP2113 en este metabolismo todavía no está completamente claro, requiriendo más esfuerzos para conocer cuál es la reacción que cataliza. Finalmente, la cepa Δ LLM-BVMO es capaz de acumular 1,2-dehidroprogesterona a partir de colesterol, pero requiere de muchas mejoras para poder ser aplicada como factoría celular en la producción de este esteroide a nivel industrial.



VIII. CONCLUSIONES

El trabajo presentado en esta Tesis Doctoral ha dado lugar a las siguientes conclusiones principales:

1. Los genes *MSMEG_6561* y *MSMEG_6563* que se encuentran en el *cluster* génico 4-HBC codifican la aldolasa msSal y la reductasa msRed, respectivamente, involucradas en el metabolismo del 4-HBC.
2. La eliminación de la aldolasa msSal permite producir AD en *M. smegmatis* a partir de esteroides con un alto rendimiento y sin contaminación de 4-HBC.
3. La sobreexpresión de los genes *MSMEG_6561* y *MSMEG_6563* incrementa la producción de 4-HBC en *M. smegmatis* a partir de esteroides.
4. Se ha completado la colección del *Golden Standard* con el *ori* RSF1010, lo que permite la aplicación de esta herramienta genética en micobacterias y la ampliación de la librería de promotores, RBS, terminadores y sistemas inducibles.
5. La expresión heteróloga del CYP450scc, la Fdx y la FdxR de *N. aromaticivorans* en *M. smegmatis* permite la producción de progesterona a partir de esteroides en una única fermentación.
6. Se ha demostrado que los genes *EGO55_13845* y *EGO55_13860* de *C. tardaugens* contenidos en el *cluster* génico *pdc* y que codifican las monooxigenasas BVMO y LLM, respectivamente, son responsables de transformar la progesterona en acetato de testosterona que, posteriormente, se degrada a través del catabolismo de la testosterona.
7. Se ha comprobado que la ruta de degradación del colesterol en *C. tardaugens* se conecta con el catabolismo de la progesterona, siendo la 1,2-dehidroprogesterona un intermediario catabólico.
8. El doble mutante Δ LLM-BVMO de *C. tardaugens* no puede metabolizar el colesterol y acumula 1,2-dehidroprogesterona a partir de colesterol, por lo que tiene un gran potencial para utilizarse como factoría celular en la producción de progestágenos a partir de esteroides.



IX. BIBLIOGRAFÍA

- Abdel-Kareem, M.M., Zohri, A.-N.A., Rasmeý, A.-H.M., Hawary, H., 2024. Enhancing the biotransformation of progesterone to the anticancer compound testololactone by *Penicillium chrysogenum* Ras3009: kinetic modelling and efficiency maximization. BMC Biotechnol. 24, 73. <https://doi.org/10.1186/s12896-024-00896-9>
- Abdi Ghavidel, A., Jajarmi, V., Bandehpour, M., Kazemi, B., 2022. Polycistronic expression of multi-subunit complexes in the eukaryotic environment: a narrative review. J. Parasitol. 17, 286–295. <https://doi.org/10.18502/ijpa.v17i3.10618>
- Agents Classified by the IARC Monographs, Volumes 1–138 – IARC Monographs on the Identification of Carcinogenic Hazards to Humans [WWW Document], 2025. Monogr. Identif. Carcinog. Hazards Hum. URL <https://monographs.iarc.who.int/agents-classified-by-the-iarc/> (accessed 4.24.25).
- Alsen, M., Sinclair, C., Cooke, P., Ziadkhanpour, K., Genden, E., van Gerwen, M., 2021. Endocrine disrupting chemicals and thyroid cancer: an overview. Toxics 9, 14. <https://doi.org/10.3390/toxics9010014>
- Anders, S., Pyl, P.T., Huber, W., 2015. HTSeq--a Python framework to work with high-throughput sequencing data. Bioinforma. Oxf. Engl. 31, 166–169. <https://doi.org/10.1093/bioinformatics/btu638>
- Arai, H., Akahira, S., Ohishi, T., Maeda, M., Kudo, T., 1998. Adaptation of *Comamonas testosteroni* TA441 to utilize phenol: organization and regulation of the genes involved in phenol degradation. Microbiol. Read. Engl. 144 (Pt 10), 2895–2903. <https://doi.org/10.1099/00221287-144-10-2895>
- Arima, K., Nagasawa, M., Bae, M., Tamura, G., 1969. Microbial transformation of sterols: part i. decomposition of cholesterol by microorganismsPart II. cleavage of sterol side chains by microorganisms. Agric. Biol. Chem. 33, 1636–1650. <https://doi.org/10.1080/00021369.1969.10859516>
- Arnold, F.M., Hohl, M., Remm, S., Koliwer-Brandl, H., Adenau, S., Chusri, S., Sander, P., Hilbi, H., Seeger, M.A., 2018. A uniform cloning platform for mycobacterial genetics and protein production. Sci. Rep. 8, 9539. <https://doi.org/10.1038/s41598-018-27687-5>
- Atienzar, F.A., Billinghamurst, Z., Depledge, M.H., 2002. 4-n-Nonylphenol and 17-beta estradiol may induce common DNA effects in developing barnacle larvae. Environ. Pollut. 120, 735–738. [https://doi.org/10.1016/s0269-7491\(02\)00184-7](https://doi.org/10.1016/s0269-7491(02)00184-7)
- Azizi, D., Arif, A., Blair, D., Dionne, J., Filion, Y., Ouarda, Y., Pazmino, A.G., Pulicharla, R., Rilstone, V., Tiwari, B., Vignale, L., Brar, S.K., Champagne, P., Drogui, P., Langlois, V.S., Blais, J.-F., 2022. A comprehensive review on current technologies for removal of endocrine disrupting chemicals from wastewaters. Environ. Res. 207, 112196. <https://doi.org/10.1016/j.envres.2021.112196>
- Balázs, A., Krifaton, C., Risa, A., Cserhádi, M., Kukolya, J., Tóth, Á., Háhn, J., Szoboszlai, S., Eldridge, M., Wang, J., Kriszt, B., 2014. Biodegradation of 5 α -dihydrotestosterone to non-androgenic products. Int. Biodeterior. Biodegrad. 93, 162–167. <https://doi.org/10.1016/j.ibiod.2014.05.008>
- Barrientos, Á., Merino, E., Casabon, I., Rodríguez, J., Crowe, A.M., Holert, J., Philipp, B., Eltis, L.D., Olivera, E.R., Luengo, J.M., 2015. Functional analyses of three acyl-CoA synthetases involved in bile acid degradation in *Pseudomonas putida* DOC21. Environ. Microbiol. 17, 47–63. <https://doi.org/10.1111/1462-2920.12395>
- Bastos, R.G., Borsuk, S., Seixas, F.K., Dellagostin, O.A., 2009. Recombinant *Mycobacterium bovis* BCG. Vaccine 27, 6495–6503. <https://doi.org/10.1016/j.vaccine.2009.08.044>

- Bergstrand, L.H., Cardenas, E., Holert, J., Van Hamme, J.D., Mohn, W.W., 2016. Delineation of steroid-degrading microorganisms through comparative genomic analysis. *mBio* 7, e00166. <https://doi.org/10.1128/mBio.00166-16>
- Bertram, R., Hillen, W., 2008. The application of Tet repressor in prokaryotic gene regulation and expression. *Microb. Biotechnol.* 1, 2–16. <https://doi.org/10.1111/j.1751-7915.2007.00001.x>
- Billinghurst, Z., Clare, A.S., Matsumura, K., Depledge, M.H., 2000. Induction of cypris major protein in barnacle larvae by exposure to 4-*n*-nonylphenol and 17 β -oestradiol. *Aquat. Toxicol.* 47, 203–212. [https://doi.org/10.1016/S0166-445X\(99\)00018-1](https://doi.org/10.1016/S0166-445X(99)00018-1)
- Bird, C.W., Lynch, J.M., Pirt, F.J., Reid, W.W., 1971. Steroids and squalene in *Methylococcus capsulatus* grown on methane. *Nature* 230, 473–474. <https://doi.org/10.1038/230473a0>
- Bisagni, S., Hatti-Kaul, R., Mamo, G., 2014. Cloning, expression and characterization of a versatile Baeyer-Villiger monooxygenase from *Dietzia* sp. D5. *AMB Express* 4, 23. <https://doi.org/10.1186/s13568-014-0023-1>
- Björkhem, I., Eggertsen, G., 2001. Genes involved in initial steps of bile acid synthesis. *Curr. Opin. Lipidol.* 12, 97–103. <https://doi.org/10.1097/00041433-200104000-00002>
- Blázquez, B., León, D.S., Torres-Bacete, J., Gómez-Luengo, Á., Kniewel, R., Martínez, I., Sordon, S., Wilczak, A., Salgado, S., Huszcza, E., Popłoński, J., Prieto, A., Nogales, J., 2023. Golden Standard: a complete standard, portable, and interoperative MoClo tool for model and non-model proteobacteria. *Nucleic Acids Res.* 51, e98. <https://doi.org/10.1093/nar/gkad758>
- Bolger, A.M., Lohse, M., Usadel, B., 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinforma.* 30, 2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>
- Bradford, M.M., 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* 72, 248–254. <https://doi.org/10.1006/abio.1976.9999>
- Brown, R.L., Peterson, G.E., 1966. Cholesterol oxidation by soil actinomycetes. *Microbiology* 45, 441–450. <https://doi.org/10.1099/00221287-45-3-441>
- Bzrostek, A., Rumijowska-Galewicz, A., Dziadek, B., Wojcik, E.A., Dziadek, J., 2013. ChoD and HsdD can be dispensable for cholesterol degradation in mycobacteria. *J. Steroid Biochem. Mol. Biol.* 134, 1–7. <https://doi.org/10.1016/j.jsbmb.2012.09.028>
- Cagnacci, A., Grandi, G., Capobianco, G., Fulghesu, A.M., Morgante, G., Biondelli, V., Piccolo, E., Casolati, E., Mangrella, M., 2025. Effects of a monophasic hormonal contraceptive with norgestimate+ethinyl estradiol on menstrual bleeding: protocol and design of a multicenter, prospective, open-label, noncomparative study in Italy. *JMIR Res. Protoc.* 14, e63683. <https://doi.org/10.2196/63683>
- Cai, H., Liu, Y., Zhang, Z., Wang, X., Song, X., Wen, Y., 2025. Enhanced removal of endocrine disrupting chemicals by chemically modified biochar: Efficiency and mechanisms. *J. Ind. Eng. Chem.* 144, 192–209. <https://doi.org/10.1016/j.jiec.2024.10.012>
- Calero, P., Nikel, P.I., 2019. Chasing bacterial chassis for metabolic engineering: a perspective review from classical to non-traditional microorganisms. *Microb. Biotechnol.* 12, 98–124. <https://doi.org/10.1111/1751-7915.13292>
- Carballeira, J.D., Quezada, M.A., Hoyos, P., Simeó, Y., Hernaiz, M.J., Alcantara, A.R., Sinisterra, J.V., 2009. Microbial cells as catalysts for stereoselective red-ox reactions. *Biotechnol. Adv.* 27, 686–714. <https://doi.org/10.1016/j.biotechadv.2009.05.001>

- Casabon, I., Zhu, S.-H., Otani, H., Liu, J., Mohn, W.W., Eltis, L.D., 2013. Regulation of the KstR2 regulon of *Mycobacterium tuberculosis* by a cholesterol catabolite. *Mol. Microbiol.* 89, 1201–1212. <https://doi.org/10.1111/mmi.12340>
- Chang, Y., Mead, D., Dhodda, V., Brumm, P., Fox, B.G., 2009. One-plasmid tunable coexpression for mycobacterial protein-protein interaction studies. *Protein Sci. Publ. Protein Soc.* 18, 2316–2325. <https://doi.org/10.1002/pro.242>
- Chen, Y.-L., Yu, C.-P., Lee, T.-H., Goh, K.-S., Chu, K.-H., Wang, P.-H., Ismail, W., Shih, C.-J., Chiang, Y.-R., 2017. Biochemical mechanisms and catabolic enzymes involved in bacterial estrogen degradation pathways. *Cell Chem. Biol.* 24, 712–724.e7. <https://doi.org/10.1016/j.chembiol.2017.05.012>
- Cheng, W., Ren, J., Huang, Q., Long, H., Jin, H., Zhang, L., Liu, H., van Ofwegen, L., Lin, W., 2016. Pregnane steroids from a gorgonian coral *Subergorgia suberosa* with anti-flu virus effects. *Steroids* 108, 99–104. <https://doi.org/10.1016/j.steroids.2016.02.003>
- Chiang, Y.-R., Ismail, W., Gallien, S., Heintz, D., Van Dorsselaer, A., Fuchs, G., 2008a. Cholest-4-en-3-one- Δ^1 -dehydrogenase, a flavoprotein catalyzing the second step in anoxic cholesterol metabolism. *Appl. Environ. Microbiol.* 74, 107–113. <https://doi.org/10.1128/AEM.01968-07>
- Chiang, Y.-R., Ismail, W., Heintz, D., Schaeffer, C., Van Dorsselaer, A., Fuchs, G., 2008b. Study of anoxic and oxic cholesterol metabolism by *Sterolibacterium denitrificans*. *J. Bacteriol.* 190, 905–914. <https://doi.org/10.1128/JB.01525-07>
- Chipley, J.R., Dreyfuss, M.S., Smucker, R.A., 1975. Cholesterol metabolism by *Mycobacterium*. *Microbios* 12, 199–207.
- Claassens, N.J., Finger-Bou, M., Scholten, B., Muis, F., de Groot, J.J., de Gier, J.-W., de Vos, W.M., van der Oost, J., 2019. Bicistronic design-based continuous and high-level membrane protein production in *Escherichia coli*. *ACS Synth. Biol.* 8, 1685–1690. <https://doi.org/10.1021/acssynbio.9b00101>
- De, D., Seth, M., Bhaduri, A.P., 1991. Regioselective reactions of 1,2-dehydroprogesterone: syntheses of pregnane derivatives as possible contragestational agents. *Steroids* 56, 189–194. [https://doi.org/10.1016/0039-128x\(91\)90081-6](https://doi.org/10.1016/0039-128x(91)90081-6)
- de Lorenzo, V., Timmis, K.N., 1994. Analysis and construction of stable phenotypes in gram-negative bacteria with Tn5- and Tn10-derived minitransposons. *Methods Enzymol.* 235, 386–405. [https://doi.org/10.1016/0076-6879\(94\)35157-0](https://doi.org/10.1016/0076-6879(94)35157-0)
- Ding, J.-Y., Shiu, J.-H., Chen, W.-M., Chiang, Y.-R., Tang, S.-L., 2016. Genomic insight into the host-endosymbiont relationship of *Endozoicomonas montiporae* CL-33(T) with its Coral Host. *Front. Microbiol.* 7, 251. <https://doi.org/10.3389/fmicb.2016.00251>
- Döhr, O., Paine, M.J.I., Friedberg, T., Roberts, G.C.K., Wolf, C.R., 2001. Engineering of a functional human NADH-dependent cytochrome P450 system. *Proc. Natl. Acad. Sci.* 98, 81–86. <https://doi.org/10.1073/pnas.98.1.81>
- Dokladda, K., Billamas, P., Palittapongarnpim, P., 2015. Different behaviours of promoters in *Mycobacterium tuberculosis* H37Rv and H37Ra. *World J. Microbiol. Biotechnol.* 31, 407–413. <https://doi.org/10.1007/s11274-014-1794-x>
- Donova, M., 2021. Microbial steroid production technologies: current trends and prospects. *Microorganisms* 10, 53. <https://doi.org/10.3390/microorganisms10010053>
- Donova, M., 2018. Microbiotechnologies for steroid production. *Microbiol. Aust.* 39, 126–129. <https://doi.org/10.1071/MA18039>

- Donova, M.V., Egorova, O.V., 2012. Microbial steroid transformations: current state and prospects. *Appl. Microbiol. Biotechnol.* 94, 1423–1447. <https://doi.org/10.1007/s00253-012-4078-0>
- Dresen, C., Lin, L.Y.-C., D'Angelo, I., Tocheva, E.I., Strynadka, N., Eltis, L.D., 2010. A flavin-dependent monooxygenase from *Mycobacterium tuberculosis* involved in cholesterol catabolism. *J. Biol. Chem.* 285, 22264–22275. <https://doi.org/10.1074/jbc.M109.099028>
- Duport, C., Spagnoli, R., Degryse, E., Pompon, D., 1998. Self-sufficient biosynthesis of pregnenolone and progesterone in engineered yeast. *Nat. Biotechnol.* 16, 186–189. <https://doi.org/10.1038/nbt0298-186>
- Efimova, V.S., Isaeva, L.V., Rubtsov, M.A., Novikova, L.A., 2019. Analysis of *in vivo* activity of the bovine cholesterol hydroxylase/lyase system proteins expressed in *Escherichia coli*. *Mol. Biotechnol.* 61, 261–273. <https://doi.org/10.1007/s12033-019-00158-6>
- Egan, S.M., Schleif, R.F., 1993. A regulatory cascade in the induction of *rhaBAD*. *J. Mol. Biol.* 234, 87–98. <https://doi.org/10.1006/jmbi.1993.1565>
- Engler, C., Youles, M., Gruetzner, R., Ehnert, T.-M., Werner, S., Jones, J.D.G., Patron, N.J., Marillonnet, S., 2014. A golden gate modular cloning toolbox for plants. *ACS Synth. Biol.* 3, 839–843. <https://doi.org/10.1021/sb4001504>
- Ewen, K.M., Ringle, M., Bernhardt, R., 2012. Adrenodoxin—A versatile ferredoxin. *IUBMB Life* 64, 506–512. <https://doi.org/10.1002/iub.1029>
- Fahrbach, M., Kuever, J., Meinke, R., Kämpfer, P., Hollender, J., 2006. *Denitratisoma oestradiolicum* gen. nov., sp. nov., a 17 β -oestradiol-degrading, denitrifying betaproteobacterium. *Int. J. Syst. Evol. Microbiol.* 56, 1547–1552. <https://doi.org/10.1099/ijs.0.63672-0>
- Fernandes, P., Cruz, A., Angelova, B., Pinheiro, H.M., Cabral, J.M.S., 2003. Microbial conversion of steroid compounds: recent developments. *Enzyme Microb. Technol.* 32, 688–705. [https://doi.org/10.1016/S0141-0229\(03\)00029-2](https://doi.org/10.1016/S0141-0229(03)00029-2)
- Fernández de Las Heras, L., García Fernández, E., María Navarro Llorens, J., Perera, J., Drzyzga, O., 2009. Morphological, physiological, and molecular characterization of a newly isolated steroid-degrading actinomycete, identified as *rhodococcus ruber* strain Chol-4. *Curr. Microbiol.* 59, 548–553. <https://doi.org/10.1007/s00284-009-9474-z>
- Fernández de Las Heras, L., Perera, J., Navarro Llorens, J.M., 2014. Cholesterol to cholestenone oxidation by ChoG, the main extracellular cholesterol oxidase of *Rhodococcus ruber* strain Chol-4. *J. Steroid Biochem. Mol. Biol.* 139, 33–44. <https://doi.org/10.1016/j.jsbmb.2013.10.001>
- Fernández de las Heras, L., van der Geize, R., Drzyzga, O., Perera, J., María Navarro Llorens, J., 2012. Molecular characterization of three 3-ketosteroid- $\Delta(1)$ -dehydrogenase isoenzymes of *Rhodococcus ruber* strain Chol-4. *J. Steroid Biochem. Mol. Biol.* 132, 271–281. <https://doi.org/10.1016/j.jsbmb.2012.06.005>
- Fernández, L., Louvado, A., Esteves, V.I., Gomes, N.C.M., Almeida, A., Cunha, Â., 2017. Biodegradation of 17 β -estradiol by bacteria isolated from deep sea sediments in aerobic and anaerobic media. *J. Hazard. Mater.* 323, 359–366. <https://doi.org/10.1016/j.jhazmat.2016.05.029>
- Fernández-Cabezón, L., Galán, B., García, J.L., 2018a. New insights on steroid biotechnology. *Front. Microbiol.* 9, 958. <https://doi.org/10.3389/fmicb.2018.00958>
- Fernández-Cabezón, L., Galán, B., García, J.L., 2018b. Unravelling a new catabolic pathway of C-19 steroids in *Mycobacterium smegmatis*. *Environ. Microbiol.* 20, 1815–1827. <https://doi.org/10.1111/1462-2920.14114>

- Ferreira, N.P., Tracey, R.P., 1984. Numerical taxonomy of cholesterol-degrading soil bacteria. *J. Appl. Bacteriol.* 57, 429–446. <https://doi.org/10.1111/j.1365-2672.1984.tb01409.x>
- Ferroni, F.M., Tolmie, C., Smit, M.S., Opperman, D.J., 2016. Structural and catalytic characterization of a fungal Baeyer-Villiger monooxygenase. *PLOS ONE* 11, e0160186. <https://doi.org/10.1371/journal.pone.0160186>
- Franceschini, S., van Beek, H.L., Pennetta, A., Martinoli, C., Fraaije, M.W., Mattevi, A., 2012. Exploring the structural basis of substrate preferences in Baeyer-Villiger monooxygenases: insight from steroid monooxygenase. *J. Biol. Chem.* 287, 22626–22634. <https://doi.org/10.1074/jbc.M112.372177>
- Frank, D.J., Waddling, C.A., La, M., Ortiz de Montellano, P.R., 2015. Cytochrome P450 125A4, the third cholesterol C-26 hydroxylase from *Mycobacterium smegmatis*. *Biochemistry* 54, 6909–6916. <https://doi.org/10.1021/acs.biochem.5b01029>
- Freier, T.A., Beitz, D.C., Li, L., Hartman, P.A., 1994. Characterization of *Eubacterium coprostanoligenes* sp. nov., a cholesterol-reducing anaerobe. *Int. J. Syst. Bacteriol.* 44, 137–142. <https://doi.org/10.1099/00207713-44-1-137>
- Frena, M., Souza, M.R.R., Damasceno, F.C., Madureira, L.A.S., Alexandre, M.R., 2016. Evaluation of anthropogenic contamination using sterol markers in a tropical estuarine system of northeast Brazil. *Mar. Pollut. Bull.* 109, 619–623. <https://doi.org/10.1016/j.marpolbul.2016.05.022>
- Frey, J., Bagdasarian, M.M., Bagdasarian, M., 1992. Replication and copy number control of the broad-host-range plasmid RSF1010. *Gene* 113, 101–106. [https://doi.org/10.1016/0378-1119\(92\)90675-f](https://doi.org/10.1016/0378-1119(92)90675-f)
- Froehner, S., Sáñez, J., 2013. Evaluation of potential sewage contamination by fecal sterol biomarkers adsorbed in natural biofilms. *Environ. Sci. Process. Impacts* 15, 2080–2086. <https://doi.org/10.1039/c3em00025g>
- Fucic, A., Gamulin, M., Ferencic, Z., Katic, J., Krayner von Krauss, M., Bartonova, A., Merlo, D.F., 2012. Environmental exposure to xenoestrogens and oestrogen related cancers: reproductive system, breast, lung, kidney, pancreas, and brain. *Environ. Health Glob. Access Sci. Source* 11 Suppl 1, S8. <https://doi.org/10.1186/1476-069X-11-S1-S8>
- Fujii, K., Satomi, M., Morita, N., Motomura, T., Tanaka, T., Kikuchi, S., 2003. *Novosphingobium tardaugens* sp. nov., an oestradiol-degrading bacterium isolated from activated sludge of a sewage treatment plant in Tokyo. *Int. J. Syst. Evol. Microbiol.* 53, 47–52. <https://doi.org/10.1099/ijs.0.02301-0>
- Galán, B., Uhía, I., García-Fernández, E., Martínez, I., Bahillo, E., de la Fuente, J.L., Barredo, J.L., Fernández-Cabezón, L., García, J.L., 2017. *Mycobacterium smegmatis* is a suitable cell factory for the production of steroidic synthons. *Microb. Biotechnol.* 10, 138–150. <https://doi.org/10.1111/1751-7915.12429>
- Gan, L., von Moltke, L.L., Trepanier, L.A., Harmatz, J.S., Greenblatt, D.J., Court, M.H., 2009. Role of NADPH-cytochrome P450 reductase and cytochrome-b5/NADH-b5 reductase in variability of CYP3A activity in human liver microsomes. *Drug Metab. Dispos.* 37, 90–96. <https://doi.org/10.1124/dmd.108.023424>
- García, J.L., Uhía, I., Galán, B., 2012. Catabolism and biotechnological applications of cholesterol degrading bacteria. *Microb. Biotechnol.* 5, 679–699. <https://doi.org/10.1111/j.1751-7915.2012.00331.x>
- García, R., Gemperlein, K., Müller, R., 2014. *Minicystis rosea* gen. nov., sp. nov., a polyunsaturated fatty acid-rich and steroid-producing soil myxobacterium. *Int. J. Syst. Evol. Microbiol.* 64, 3733–3742. <https://doi.org/10.1099/ijs.0.068270-0>

- García-Fernández, E., Frank, D.J., Galán, B., Kells, P.M., Podust, L.M., García, J.L., Ortiz de Montellano, P.R., 2013. A highly conserved mycobacterial cholesterol catabolic pathway. *Environ. Microbiol.* 15, 2342–2359. <https://doi.org/10.1111/1462-2920.12108>
- García-Fernández, E., Medrano, F.J., Galán, B., García, J.L., 2014. Deciphering the transcriptional regulation of cholesterol catabolic pathway in mycobacteria: identification of the inducer of KstR repressor. *J. Biol. Chem.* 289, 17576–17588. <https://doi.org/10.1074/jbc.M113.545715>
- García-Fernández, J., Galán, B., Felpeto-Santero, C., Barredo, J.L., García, J.L., 2017a. Production of 4-Ene-3-ketosteroids in *Corynebacterium glutamicum*. *Catalysts* 7, 316. <https://doi.org/10.3390/catal7110316>
- García-Fernández, J., Galán, B., Medrano, F.J., García, J.L., 2015. Characterization of the KstR2 regulator responsible of the lower cholesterol degradative pathway in *Mycobacterium smegmatis*. *Environ. Microbiol. Rep.* 7, 155–163. <https://doi.org/10.1111/1758-2229.12255>
- García-Fernández, J., Papavinasasundaram, K., Galán, B., Sassetti, C.M., García, J.L., 2017b. Molecular and functional analysis of the mce4 operon in *Mycobacterium smegmatis*. *Environ. Microbiol.* 19, 3689–3699. <https://doi.org/10.1111/1462-2920.13869>
- Gawas, D., Garcia, R., Huch, V., Müller, R., 2011. A Highly Conjugated Dihydroxylated C28 Steroid from a *Myxobacterium*. *J. Nat. Prod.* 74, 1281–1283. <https://doi.org/10.1021/np100682c>
- Gawin, A., Valla, S., Brautaset, T., 2017. The XylS/Pm regulator/promoter system and its use in fundamental studies of bacterial gene expression, recombinant protein production and metabolic engineering. *Microb. Biotechnol.* 10, 702–718. <https://doi.org/10.1111/1751-7915.12701>
- Gerber, A., Kleser, M., Biedendieck, R., Bernhardt, R., Hannemann, F., 2015. Functionalized PHB granules provide the basis for the efficient side-chain cleavage of cholesterol and analogs in recombinant *Bacillus megaterium*. *Microb. Cell Factories* 14, 107. <https://doi.org/10.1186/s12934-015-0300-y>
- Giró, M., Carrillo, N., Krapp, A.R., 2006. Glucose-6-phosphate dehydrogenase and ferredoxin-NADP(H) reductase contribute to damage repair during the soxRS response of *Escherichia coli*. *Microbiology* 152, 1119–1128. <https://doi.org/10.1099/mic.0.28612-0>
- Godoy, M.S., de Miguel, S.R., Prieto, M.A., 2023. A singular PpaA/AerR-like protein in *Rhodospirillum rubrum* rules beyond the boundaries of photosynthesis in response to the intracellular redox state. *mSystems* 8, e0070223. <https://doi.org/10.1128/msystems.00702-23>
- Golovko, O., Šauer, P., Fedorova, G., Kroupová, H.K., Grabic, R., 2018. Determination of progestogens in surface and waste water using SPE extraction and LC-APCI/APPI-HRPS. *Sci. Total Environ.* 621, 1066–1073. <https://doi.org/10.1016/j.scitotenv.2017.10.120>
- Gong, W., Kisiela, M., Schilhabel, M.B., Xiong, G., Maser, E., 2012. Genome sequence of *Comamonas testosteroni* ATCC 11996, a representative strain involved in steroid degradation. *J. Bacteriol.* 194, 1633–1634. <https://doi.org/10.1128/JB.06795-11>
- Grant, S.G., Jessee, J., Bloom, F.R., Hanahan, D., 1990. Differential plasmid rescue from transgenic mouse DNAs into *Escherichia coli* methylation-restriction mutants. *Proc. Natl. Acad. Sci. U. S. A.* 87, 4645–4649. <https://doi.org/10.1073/pnas.87.12.4645>
- Griffin, J.E., Pandey, A.K., Gilmore, S.A., Mizrahi, V., McKinney, J.D., Bertozzi, C.R., Sassetti, C.M., 2012. Cholesterol catabolism by *Mycobacterium tuberculosis* requires transcriptional and metabolic adaptations. *Chem. Biol.* 19, 218–227. <https://doi.org/10.1016/j.chembiol.2011.12.016>
- Griffith, D.R., Kido Soule, M.C., Eglinton, T.I., Kujawinski, E.B., Gschwend, P.M., 2016. Steroidal estrogen sources in a sewage-impacted coastal ocean. *Environ. Sci. Process. Impacts* 18, 981–991. <https://doi.org/10.1039/c6em00127k>

- Han, S., Liu, X., He, B., Zhai, X., Yuan, C., Li, Y., Lin, W., Wang, H., Zhang, B., 2024. Efficient production of 9,22-dihydroxy-23,24-bisnorchol-4-ene-3-one from phytosterols by modifying multiple genes in *Mycobacterium fortuitum*. *Int. J. Mol. Sci.* 25, 3579. <https://doi.org/10.3390/ijms25073579>
- Hanselman, T.A., Graetz, D.A., Wilkie, A.C., 2003. Manure-borne estrogens as potential environmental contaminants: a review. *Environ. Sci. Technol.* 37, 5471–5478. <https://doi.org/10.1021/es034410+>
- Hanson, J.R., 2005. Steroids: reactions and partial synthesis. *Nat. Prod. Rep.* 22, 104–110. <https://doi.org/10.1039/B415086B>
- Harder, J., Probian, C., 1997. Anaerobic mineralization of cholesterol by a novel type of denitrifying bacterium. *Arch. Microbiol.* 167, 269–274. <https://doi.org/10.1007/s002030050442>
- Harrault, L., Milek, K., Jardé, E., Jeanneau, L., Derrien, M., Anderson, D.G., 2019. Faecal biomarkers can distinguish specific mammalian species in modern and past environments. *PLOS ONE* 14, e0211119. <https://doi.org/10.1371/journal.pone.0211119>
- He, D., Zhu, T., Sun, J., Pan, X., Li, J., Luo, H., 2024. Emerging organic contaminants in sewage sludge: Current status, technological challenges and regulatory perspectives. *Sci. Total Environ.* 955, 177234. <https://doi.org/10.1016/j.scitotenv.2024.177234>
- Hernández-Fernández, G., Acedos, M.G., de la Torre, I., Ibero, J., García, J.L., Galán, B., 2024. Improving the production of 22-hydroxy-23,24-bisnorchol-4-ene-3-one in *Mycolicibacterium smegmatis*. *Microb. Biotechnol.* 17, e14551. <https://doi.org/10.1111/1751-7915.14551>
- Hernández-Fernández, G., Acedos, M.G., García, J.L., Galán, B., 2023. Identification of the aldolase responsible for the production of 22-hydroxy-23,24-bisnorchol-4-ene-3-one from natural sterols in *Mycolicibacterium smegmatis*. *Microb. Biotechnol.* 17, e14270. <https://doi.org/10.1111/1751-7915.14270>
- Ho, N.A.T., Dawes, S.S., Crowe, A.M., Casabon, I., Gao, C., Kendall, S.L., Baker, E.N., Eltis, L.D., Lott, J.S., 2016. The structure of the transcriptional repressor KstR in complex with CoA thioester cholesterol metabolites sheds light on the regulation of cholesterol catabolism in *Mycobacterium tuberculosis*. *J. Biol. Chem.* 291, 7256–7266. <https://doi.org/10.1074/jbc.M115.707760>
- Holert, J., Yücel, O., Jagmann, N., Prestel, A., Möller, H.M., Philipp, B., 2016. Identification of bypass reactions leading to the formation of one central steroid degradation intermediate in metabolism of different bile salts in *Pseudomonas* sp. strain Chol1. *Environ. Microbiol.* 18, 3373–3389. <https://doi.org/10.1111/1462-2920.13192>
- Horinouchi, M., Hayashi, T., Koshino, H., Kurita, T., Kudo, T., 2005. Identification of 9,17-dioxo-1,2,3,4,10,19-hexanorandrostane-5-oic acid, 4-hydroxy-2-oxohexanoic acid, and 2-hydroxyhexa-2,4-dienoic acid and related enzymes involved in testosterone degradation in *Comamonas testosteroni* TA441. *Appl. Environ. Microbiol.* 71, 5275–5281. <https://doi.org/10.1128/AEM.71.9.5275-5281.2005>
- Horinouchi, M., Hayashi, T., Kudo, T., 2012. Steroid degradation in *Comamonas testosteroni*. *J. Steroid Biochem. Mol. Biol.* 129, 4–14. <https://doi.org/10.1016/j.jsbmb.2010.10.008>
- Horinouchi, M., Hayashi, T., Kudo, T., 2004a. The genes encoding the hydroxylase of 3-hydroxy-9,10-secoandrostane-1,3,5(10)-triene-9,17-dione in steroid degradation in *Comamonas testosteroni* TA441. *J. Steroid Biochem. Mol. Biol.* 92, 143–154. <https://doi.org/10.1016/j.jsbmb.2004.09.002>
- Horinouchi, M., Hayashi, T., Yamamoto, T., Kudo, T., 2003. A new bacterial steroid degradation gene cluster in *Comamonas testosteroni* TA441 which consists of aromatic-compound degradation genes for seco-steroids and 3-ketosteroid dehydrogenase genes. *Appl. Environ. Microbiol.* 69, 4421–4430. <https://doi.org/10.1128/AEM.69.8.4421-4430.2003>

- Horinouchi, M., Koshino, H., Malon, M., Hirota, H., Hayashi, T., 2019. Steroid degradation in *Comamonas testosteroni* TA441: identification of the entire β -oxidation cycle of the cleaved B ring. *Appl. Environ. Microbiol.* 85, e01204-19. <https://doi.org/10.1128/AEM.01204-19>
- Horinouchi, M., Kurita, T., Hayashi, T., Kudo, T., 2010. Steroid degradation genes in *Comamonas testosteroni* TA441: Isolation of genes encoding a $\Delta 4(5)$ -isomerase and 3α - and 3β -dehydrogenases and evidence for a 100kb steroid degradation gene hot spot. *J. Steroid Biochem. Mol. Biol.* 122, 253–263. <https://doi.org/10.1016/j.jsbmb.2010.06.002>
- Horinouchi, M., Kurita, T., Yamamoto, T., Hatori, E., Hayashi, T., Kudo, T., 2004b. Steroid degradation gene cluster of *Comamonas testosteroni* consisting of 18 putative genes from meta-cleavage enzyme gene *tesB* to regulator gene *tesR*. *Biochem. Biophys. Res. Commun.* 324, 597–604. <https://doi.org/10.1016/j.bbrc.2004.09.096>
- Horinouchi, M., Yamamoto, T., Taguchi, K., Arai, H., Kudo, T., 2001. Meta-cleavage enzyme gene *tesB* is necessary for testosterone degradation in *Comamonas testosteroni* TA441. *Microbiol. Read. Engl.* 147, 3367–3375. <https://doi.org/10.1099/00221287-147-12-3367>
- Ibero, J., 2022. Estudio de la degradación de esteroides en *Novosphingobium tardaugens* NBRC 16725 y sus aplicaciones biotecnológicas. Universidad Complutense de Madrid, Madrid.
- Ibero, J., Galán, B., Díaz, E., García, J.L., 2019. Testosterone degradative pathway of *Novosphingobium tardaugens*. *Genes* 10, 871. <https://doi.org/10.3390/genes10110871>
- Ibero, J., Galán, B., García, J.L., 2021. Identification of the EdcR estrogen-dependent repressor in *Caenibius tardaugens* NBRC 16725: construction of a cellular estradiol biosensor. *Genes* 12, 1846. <https://doi.org/10.3390/genes12121846>
- Ibero, J., Galán, B., Rivero-Buceta, V., García, J.L., 2020. Unraveling the 17β -estradiol degradation pathway in *Novosphingobium tardaugens* NBRC 16725. *Front. Microbiol.* 11, 588300. <https://doi.org/10.3389/fmicb.2020.588300>
- Isupov, M.N., Schröder, E., Gibson, R.P., Beecher, J., Donadio, G., Saneei, V., Dcunha, S.A., McGhie, E.J., Sayer, C., Davenport, C.F., Lau, P.C.K., Hasegawa, Y., Iwaki, H., Kadow, M., Balke, K., Bornscheuer, U.T., Bourenkov, G., Littlechild, J.A., 2018. The oxygenating constituent of 3,6-diketocamphane monooxygenase from the CAM plasmid of *Pseudomonas putida*: the first crystal structure of a type II Baeyer-Villiger monooxygenase. *Corrigendum. Acta Crystallogr. Sect. Struct. Biol.* 74, 379. <https://doi.org/10.1107/S205979831800150X>
- Iverson, S.V., Haddock, T.L., Beal, J., Densmore, D.M., 2016. CIDAR MoClo: improved MoClo assembly standard and new *E. coli* part library enable rapid combinatorial design for synthetic and traditional biology. *ACS Synth. Biol.* 5, 99–103. <https://doi.org/10.1021/acssynbio.5b00124>
- Iwaki, H., Grosse, S., Bergeron, H., Leisch, H., Morley, K., Hasegawa, Y., Lau, P.C.K., 2013. Camphor pathway redux: functional recombinant expression of 2,5- and 3,6-diketocamphane monooxygenases of *Pseudomonas putida* ATCC 17453 with their cognate flavin reductase catalyzing Baeyer-Villiger reactions. *Appl. Environ. Microbiol.* 79, 3282–3293. <https://doi.org/10.1128/AEM.03958-12>
- Jackson, M., Reinaldo Camacho, L., Gicquel, B., Guilhot, C., 2001. Gene replacement and transposon delivery using the negative selection marker *sacB*. *Methods Mol. Med.* 54, 59–75. <https://doi.org/10.1385/1-59259-147-7:059>
- Jacoby, C., Ferlino, S., Bezold, D., Jessen, H., Müller, M., Boll, M., 2020. ATP-dependent hydroxylation of an unactivated primary carbon with water. *Nat. Commun.* 11, 3906. <https://doi.org/10.1038/s41467-020-17675-7>
- Janocha, S., Bichet, A., Zöllner, A., Bernhardt, R., 2011. Substitution of lysine with glutamic acid at position 193 in bovine CYP11A1 significantly affects protein oligomerization and solubility

- but not enzymatic activity. *Biochim. Biophys. Acta* 1814, 126–131. <https://doi.org/10.1016/j.bbapap.2010.06.002>
- Jarošová, B., Erseková, A., Hilscherová, K., Loos, R., Gawlik, B.M., Giesy, J.P., Bláha, L., 2014. Europe-wide survey of estrogenicity in wastewater treatment plant effluents: the need for the effect-based monitoring. *Environ. Sci. Pollut. Res.* 21, 10970–10982. <https://doi.org/10.1007/s11356-014-3056-8>
- Javid, M., Nickavar, B., Vahidi, H., Faramarzi, M.A., 2018. Baeyer-Villiger oxidation of progesterone by *Aspergillus sojae* PTCC 5196. *Steroids* 140, 52–57. <https://doi.org/10.1016/j.steroids.2018.07.008>
- Johnson, A.C., Belfroid, A., Di Corcia, A.D., 2000. Estimating steroid oestrogen inputs into activated sludge treatment works and observations on their removal from the effluent. *Sci. Total Environ.* 256, 163–173. [https://doi.org/10.1016/S0048-9697\(00\)00481-2](https://doi.org/10.1016/S0048-9697(00)00481-2)
- Kanno, A.I., Goulart, C., Rofatto, H.K., Oliveira, S.C., Leite, L.C.C., McFadden, J., 2016. New recombinant *Mycobacterium bovis* BCG expression vectors: improving genetic control over mycobacterial promoters. *Appl. Environ. Microbiol.* 82, 2240–2246. <https://doi.org/10.1128/AEM.03677-15>
- Karpov, M., Strizhov, N., Novikova, L., Lobastova, T., Khomutov, S., Shutov, A., Kazantsev, A., Donova, M., 2024. Pregnenolone and progesterone production from natural sterols using recombinant strain of *Mycolicibacterium smegmatis* mc2 155 expressing mammalian steroidogenesis system. *Microb. Cell Factories* 23, 105. <https://doi.org/10.1186/s12934-024-02385-2>
- Kasambala, H.R., Rwiza, M.J., Mdegela, R.H., 2019. Levels and distribution of progesterone in receiving waters and wastewaters of a growing urban area. *Water Sci. Technol.* 80, 1107–1117. <https://doi.org/10.2166/wst.2019.350>
- Kelly, C.L., Liu, Z., Yoshihara, A., Jenkinson, S.F., Wormald, M.R., Otero, J., Estévez, A., Kato, A., Marqvorsen, M.H.S., Fleet, G.W.J., Estévez, R.J., Izumori, K., Heap, J.T., 2016. synthetic chemical inducers and genetic decoupling enable orthogonal control of the *rhaBAD* promoter. *ACS Synth. Biol.* 5, 1136–1145. <https://doi.org/10.1021/acssynbio.6b00030>
- Kendall, S.L., Withers, M., Soffair, C.N., Moreland, N.J., Gurcha, S., Sidders, B., Frita, R., Ten Bokum, A., Besra, G.S., Lott, J.S., Stoker, N.G., 2007. A highly conserved transcriptional repressor controls a large regulon involved in lipid degradation in *Mycobacterium smegmatis* and *Mycobacterium tuberculosis*. *Mol. Microbiol.* 65, 684–699. <https://doi.org/10.1111/j.1365-2958.2007.05827.x>
- Kidd, K.A., Blanchfield, P.J., Mills, K.H., Palace, V.P., Evans, R.E., Lazorchak, J.M., Flick, R.W., 2007. Collapse of a fish population after exposure to a synthetic estrogen. *Proc. Natl. Acad. Sci. U. S. A.* 104, 8897–8901. <https://doi.org/10.1073/pnas.0609568104>
- Kieslich, K., 1985. Microbial side-chain degradation of sterols. *J. Basic Microbiol.* 25, 461–474. <https://doi.org/10.1002/jobm.3620250713>
- Koga, H., Yamaguchi, E., Matsunaga, K., Aramaki, H., Horiuchi, T., 1989. Cloning and nucleotide sequences of NADH-putidaredoxin reductase gene (*camA*) and putidaredoxin gene (*camB*) involved in cytochrome P-450cam hydroxylase of *Pseudomonas putida*. *J. Biochem.* 106, 831–836. <https://doi.org/10.1093/oxfordjournals.jbchem.a122939>
- Kohl, W., Gloe, A., Reichenbach, H., 1983. Steroids from the myxobacterium *Nannocystis exedens*. *Microbiology* 129, 1629–1635. <https://doi.org/10.1099/00221287-129-6-1629>
- Kolatorova, L., Vitku, J., Suchopar, J., Hill, M., Parizek, A., 2022. Progesterone: a steroid with wide range of effects in physiology as well as human medicine. *Int. J. Mol. Sci.* 23, 7989. <https://doi.org/10.3390/ijms23147989>

- Kornman, K.S., Loesche, W.J., 1982. Effects of estradiol and progesterone on *Bacteroides melaninogenicus* and *Bacteroides gingivalis*. *Infect. Immun.* 35, 256–263. <https://doi.org/10.1128/iai.35.1.256-263.1982>
- Kurisu, F., Ogura, M., Saitoh, S., Yamazoe, A., Yagi, O., 2010. Degradation of natural estrogen and identification of the metabolites produced by soil isolates of *Rhodococcus* sp. and *Sphingomonas* sp. *J. Biosci. Bioeng.* 109, 576–582. <https://doi.org/10.1016/j.jbiosc.2009.11.006>
- Kutney, J.P., Herrington, E.J., Spassov, G., 2003. Process for fermentation of phytosterols to androstadienedione. WO2003064674A2.
- Lack, N., Lowe, E.D., Liu, J., Eltis, L.D., Noble, M.E.M., Sim, E., Westwood, I.M., 2008. Structure of HsaD, a steroid-degrading hydrolase, from *Mycobacterium tuberculosis*. *Acta Crystallograph. Sect. F Struct. Biol. Cryst. Commun.* 64, 2–7. <https://doi.org/10.1107/S1744309107065931>
- Lack, N.A., Yam, K.C., Lowe, E.D., Horsman, G.P., Owen, R.L., Sim, E., Eltis, L.D., 2010. Characterization of a carbon-carbon hydrolase from *Mycobacterium tuberculosis* involved in cholesterol metabolism. *J. Biol. Chem.* 285, 434–443. <https://doi.org/10.1074/jbc.M109.058081>
- Laemmli, U.K., 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227, 680–685. <https://doi.org/10.1038/227680a0>
- Lamb, D.C., Jackson, C.J., Warrilow, A.G.S., Manning, N.J., Kelly, D.E., Kelly, S.L., 2007. Lanosterol biosynthesis in the prokaryote *Methylococcus capsulatus*: insight into the evolution of sterol biosynthesis. *Mol. Biol. Evol.* 24, 1714–1721. <https://doi.org/10.1093/molbev/msm090>
- Lange, I.G., Daxenberger, A., Schiffer, B., Witters, H., Ibarreta, D., Meyer, H.H.D., 2002. Sex hormones originating from different livestock production systems: fate and potential disrupting activity in the environment. *Anal. Chim.* 473, 27–37. [https://doi.org/10.1016/S0003-2670\(02\)00748-1](https://doi.org/10.1016/S0003-2670(02)00748-1)
- Laveaga, G.S., 2005. Uncommon trajectories: steroid hormones, Mexican peasants, and the search for a wild yam. *Stud. Hist. Philos. Biol. Biomed. Sci.* 36, 743–760. <https://doi.org/10.1016/j.shpsc.2005.09.007>
- Li, D.-N., Liu, X.-Y., Xu, J.-B., Shi, K., Li, J.-M., Diao, N.-C., Zong, Y., Zeng, F.-L., Du, R., 2024. *Mycobacterium tuberculosis* Rv1048c affects the biological characteristics of recombinant *Mycobacterium smegmatis*. *Sci. Rep.* 14, 29749. <https://doi.org/10.1038/s41598-024-81405-y>
- Li, L., Buhman, K.K., Hartman, P.A., Beitz, D.C., 1995. Hypocholesterolemic effect of *Eubacterium coprostanoligenes* ATCC 51222 in rabbits. *Lett. Appl. Microbiol.* 20, 137–140. <https://doi.org/10.1111/j.1472-765x.1995.tb00410.x>
- Li, S., Liu, J., Williams, M.A., Ling, W., Sun, K., Lu, C., Gao, Y., Waigi, M.G., 2020. Metabolism of 17 β -estradiol by *Novosphingobium* sp. ES2-1 as probed via HRMS combined with 13C3-labeling. *J. Hazard. Mater.* 389, 121875. <https://doi.org/10.1016/j.jhazmat.2019.121875>
- Li, Wenjing, Ge, F., Zhang, Q., Ren, Y., Yuan, J., He, J., Li, Wei, Chen, G., Zhang, G., Zhuang, Y., Xu, L., 2014. Identification of gene expression profiles in the actinomycete *Gordonia neofelifaecis* grown with different steroids. *Genome* 57, 345–353. <https://doi.org/10.1139/gen-2014-0030>
- Li, X., Zheng, W., Kelly, W.R., 2013. Occurrence and removal of pharmaceutical and hormone contaminants in rural wastewater treatment lagoons. *Sci. Total Environ.* 445–446, 22–28. <https://doi.org/10.1016/j.scitotenv.2012.12.035>
- Lin, C.-W., Wang, P.-H., Ismail, W., Tsai, Y.-W., El Naya, A., Yang, C.-Y., Yang, F.-C., Wang, C.-H., Chiang, Y.-R., 2015. Substrate uptake and subcellular compartmentation of anoxic

- cholesterol catabolism in *Sterolibacterium denitrificans*. J. Biol. Chem. 290, 1155–1169. <https://doi.org/10.1074/jbc.M114.603779>
- Liu, Kun, Wang, F.-Q., Liu, Ke, Zhao, Y., Gao, B., Tao, X., Wei, D., 2022. Light-driven progesterone production by InP-(*M. neoaurum*) biohybrid system. Bioresour. Bioprocess. 9, 93. <https://doi.org/10.1186/s40643-022-00575-7>
- Liu, S., Ying, G.-G., Liu, Y.-S., Peng, F.-Q., He, L.-Y., 2013. Degradation of norgestrel by bacteria from activated sludge: comparison to progesterone. Environ. Sci. Technol. 130829113920003. <https://doi.org/10.1021/es304688g>
- Liu, S.-S., Chen, J., Zhang, J.-N., Liu, Y.-S., Hu, L.-X., Chen, X.-W., Liu, S., Xu, X.-R., Ying, G.-G., 2020. Microbial transformation of progesterone and dydrogesterone by bacteria from swine wastewater: Degradation kinetics and products identification. Sci. Total Environ. 701, 134930. <https://doi.org/10.1016/j.scitotenv.2019.134930>
- Liu, X., He, B., Zhang, J., Yuan, C., Han, S., Du, G., Shi, J., Sun, J., Zhang, B., 2023a. Phytosterol conversion into C9 non-hydroxylated derivatives through gene regulation in *Mycobacterium fortuitum*. Appl. Microbiol. Biotechnol. 107, 7635–7646. <https://doi.org/10.1007/s00253-023-12812-w>
- Liu, X., Zhang, J., Yuan, C., Du, G., Han, S., Shi, J., Sun, J., Zhang, B., 2023b. Improving the production of 9 α -hydroxy-4-androstene-3,17-dione from phytosterols by 3-ketosteroid- Δ 1-dehydrogenase deletions and multiple genetic modifications in *Mycobacterium fortuitum*. Microb. Cell Factories 22, 53. <https://doi.org/10.1186/s12934-023-02052-y>
- Lorenzen, A., Hendel, J.G., Conn, K.L., Bittman, S., Kwabiah, A.B., Lazarovitz, G., Massé, D., McAllister, T.A., Topp, E., 2004. Survey of hormone activities in municipal biosolids and animal manures. Environ. Toxicol. 19, 216–225. <https://doi.org/10.1002/tox.20014>
- Maher, C., Hassan, K.A., 2023. The Gram-negative permeability barrier: tipping the balance of the in and the out. mBio 14, e01205-23. <https://doi.org/10.1128/mbio.01205-23>
- Makeeva, D.S., Dovbnya, D.V., Donova, M.V., Novikova, L.A., 2013. Functional reconstruction of bovine P450scc steroidogenic system in *Escherichia coli*. Am. J. Mol. Biol. 03, 173–182. <https://doi.org/10.4236/ajmb.2013.34023>
- Malaviya, A., Gomes, J., 2008. Androstenedione production by biotransformation of phytosterols. Bioresour. Technol. 99, 6725–6737. <https://doi.org/10.1016/j.biortech.2008.01.039>
- Martin, C.K.A., 1977. Microbial cleavage of sterol side chains. Advances in Applied Microbiology. 29–58. [https://doi.org/10.1016/S0065-2164\(08\)70159-X](https://doi.org/10.1016/S0065-2164(08)70159-X)
- Mast, N., Annalora, A.J., Lodowski, D.T., Palczewski, K., Stout, C.D., Pikuleva, I.A., 2011. Structural basis for three-step sequential catalysis by the cholesterol side chain cleavage enzyme CYP11A1. J. Biol. Chem. 286, 5607–5613. <https://doi.org/10.1074/jbc.M110.188433>
- Miller, W.L., 1988. Molecular biology of steroid hormone synthesis. Endocr. Rev. 9, 295–318. <https://doi.org/10.1210/edrv-9-3-295>
- Mitchell, S.L., Kearns, D.B., Carlson, E.E., 2024. Penicillin-binding protein redundancy in *Bacillus subtilis* enables growth during alkaline shock. Appl. Environ. Microbiol. 90, e0054823. <https://doi.org/10.1128/aem.00548-23>
- Miyamoto, M., Matsumoto, J., Iwaya, T., Itagaki, E., 1995. Bacterial steroid monooxygenase catalyzing the Baeyer-Villiger oxidation of C21-ketosteroids from *Rhodococcus rhodochrous*: the isolation and characterization. Biochim. Biophys. Acta BBA - Protein Struct. Mol. Enzymol. 1251, 115–124. [https://doi.org/10.1016/0167-4838\(95\)00090-H](https://doi.org/10.1016/0167-4838(95)00090-H)

- Mohn, W.W., van der Geize, R., Stewart, G.R., Okamoto, S., Liu, J., Dijkhuizen, L., Eltis, L.D., 2008. The actinobacterial *mce4* locus encodes a steroid transporter. *J. Biol. Chem.* 283, 35368–35374. <https://doi.org/10.1074/jbc.M805496200>
- Moreno, A., Quereda-Moraleda, I., Lozano-Vallhonrat, C., Buñuel-Escudero, M., Botha, S., Kupitz, C., Lisova, S., Sierra, R., Mariani, V., Schleissner, P., Gee, L.B., Dörner, K., Schmidt, C., Han, H., Kloos, M., Smyth, P., Valerio, J., Schulz, J., de Wijn, R., Melo, D.V.M., Round, A., Trost, F., Sobolev, E., Juncheng, E., Sikorski, M., Bean, R., Martínez-Júlvez, M., Martín-García, J.M., Medina, M., 2024. New insights into the function and molecular mechanisms of Ferredoxin-NADP⁺ reductase from *Brucella ovis*. *Arch. Biochem. Biophys.* 762, 110204. <https://doi.org/10.1016/j.abb.2024.110204>
- Morii, S., Sawamoto, S., Yamauchi, Y., Miyamoto, M., Iwami, M., Itagaki, E., 1999. Steroid monooxygenase of *Rhodococcus rhodochrous*: sequencing of the genomic DNA, and hyperexpression, purification, and characterization of the recombinant enzyme. *J. Biochem.* 126, 624–631. <https://doi.org/10.1093/oxfordjournals.jbchem.a022494>
- Morisaki, M., Duque, C., Takane, K., Ikekawa, N., Shikita, M., 1982. Substrate specificity of adrenocortical cytochrome P-450SCC. II. Effect of structural modification of cholesterol A/B ring on their side chain cleavage reaction. *J. Steroid Biochem.* 16, 101–105. [https://doi.org/10.1016/0022-4731\(82\)90150-9](https://doi.org/10.1016/0022-4731(82)90150-9)
- Moukhtari, F.E., Martín-Pozo, L., Zafra-Gómez, A., 2023. Strategies based on the use of microorganisms for the elimination of pollutants with endocrine-disrupting activity in the environment. *J. Environ. Chem. Eng.* 11, 109268. <https://doi.org/10.1016/j.jece.2023.109268>
- Mutalik, V.K., Guimaraes, J.C., Cambray, G., Lam, C., Christoffersen, M.J., Mai, Q.-A., Tran, A.B., Paull, M., Keasling, J.D., Arkin, A.P., Endy, D., 2013. Precise and reliable gene expression via standard transcription and translation initiation elements. *Nat. Methods* 10, 354–360. <https://doi.org/10.1038/nmeth.2404>
- Nelson, D.L., Cox, M.M., Nelson, D.L., 2013. *Lehninger principles of biochemistry*, Sixth edition. ed. Macmillan Higher Education, Basingstoke.
- Nesbitt, N.M., Yang, X., Fontán, P., Kolesnikova, I., Smith, I., Sampson, N.S., Dubnau, E., 2010. A thiolase of *Mycobacterium tuberculosis* is required for virulence and production of androstenedione and androstadienedione from cholesterol. *Infect. Immun.* 78, 275–282. <https://doi.org/10.1128/IAI.00893-09>
- Oliveira, T.L., Stedman, A., Rizzi, C., Dorneles, J., da Cunha, C.E.P., Junior, A.S.V., Dellagostin, O.A., McFadden, J., 2019. A standardized BioBrick toolbox for the assembly of sequences in mycobacteria. *Tuberc. Edinb. Scotl.* 119, 101851. <https://doi.org/10.1016/j.tube.2019.07.002>
- Olivera, E.R., Luengo, J.M., 2019. Steroids as environmental compounds recalcitrant to degradation: genetic mechanisms of bacterial biodegradation pathways. *Genes* 10, 512. <https://doi.org/10.3390/genes10070512>
- Ouellet, H., Guan, S., Johnston, J.B., Chow, E.D., Kells, P.M., Burlingame, A.L., Cox, J.S., Podust, L.M., de Montellano, P.R.O., 2010. *Mycobacterium tuberculosis* CYP125A1, a steroid C27 monooxygenase that detoxifies intracellularly generated cholest-4-en-3-one. *Mol. Microbiol.* 77, 730–742. <https://doi.org/10.1111/j.1365-2958.2010.07243.x>
- Pal, N., Sharma, P., Singh, S., Ojha, R., Kumawat, M., Shubham, S., Verma, V., Tiwari, R.R., Sarma, D.K., Kumar, M., 2025. Detection and risk assessment of bisphenol-A and phthalate esters in bottled water: implications for public health. *Int. J. Environ. Health Res.* 1–11. <https://doi.org/10.1080/09603123.2025.2465855>

- Pan, H., Chang, S., Qu, Y., Liu, M., Tian, W., Chang, Z., 2023. Hydrocortisone production using whole-cell biocatalysts in recombinant *Escherichia coli*. *Biochem. Eng. J.* 198, 109023. <https://doi.org/10.1016/j.bej.2023.109023>
- Pandey, A.K., Sassetti, C.M., 2008. Mycobacterial persistence requires the utilization of host cholesterol. *Proc. Natl. Acad. Sci.* 105, 4376–4380. <https://doi.org/10.1073/pnas.0711159105>
- Parish, T., Stoker, N.G., 1998. Electroporation of mycobacteria. *Methods Mol. Biol.* Clifton NJ 101, 129–144. <https://doi.org/10.1385/0-89603-471-2:129>
- Parsi, Z., Górecki, T., 2006. Determination of ergosterol as an indicator of fungal biomass in various samples using non-discriminating flash pyrolysis. *J. Chromatogr. A* 1130, 145–150. <https://doi.org/10.1016/j.chroma.2006.07.045>
- Payne, D.W., Talalay, P., 1985. Isolation of novel microbial 3 alpha-, 3 beta-, and 17 beta-hydroxysteroid dehydrogenases. Purification, characterization, and analytical applications of a 17 beta-hydroxysteroid dehydrogenase from an *Alcaligenes* sp. *J. Biol. Chem.* 260, 13648–13655.
- Pearson, A., Budin, M., Brocks, J.J., 2003. Phylogenetic and biochemical evidence for sterol synthesis in the bacterium *Gemmata obscuriglobus*. *Proc. Natl. Acad. Sci.* 100, 15352–15357. <https://doi.org/10.1073/pnas.2536559100>
- Peng, H., Wang, Y., Jiang, K., Chen, X., Zhang, W., Zhang, Y., Deng, Z., Qu, X., 2021. A dual role reductase from phytosterols catabolism enables the efficient production of valuable steroid precursors. *Angew. Chem.* 133, 5474–5480. <https://doi.org/10.1002/ange.202015462>
- Petrusma, M., Hessels, G., Dijkhuizen, L., van der Geize, R., 2011. Multiplicity of 3-ketosteroid-9 α -hydroxylase enzymes in *Rhodococcus rhodochrous* DSM43269 for specific degradation of different classes of steroids. *J. Bacteriol.* 193, 3931–3940. <https://doi.org/10.1128/JB.00274-11>
- Piironen, V., Lindsay, D.G., Miettinen, T.A., Toivo, J., Lampi, A.-M., 2000. Plant sterols: biosynthesis, biological function and their importance to human nutrition. *J. Sci. Food Agric.* 80, 939–966. [https://doi.org/10.1002/\(SICI\)1097-0010\(20000515\)80:7<939::AID-JSFA644>3.0.CO;2-C](https://doi.org/10.1002/(SICI)1097-0010(20000515)80:7<939::AID-JSFA644>3.0.CO;2-C)
- Plésiat, P., Nikaido, H., 1992. Outer membranes of gram-negative bacteria are permeable to steroid probes. *Mol. Microbiol.* 6, 1323–1333. <https://doi.org/10.1111/j.1365-2958.1992.tb00853.x>
- Pruneda-Paz, J.L., Linares, M., Cabrera, J.E., Genti-Raimondi, S., 2004. TeiR, a LuxR-type transcription factor required for testosterone degradation in *Comamonas testosteroni*. *J. Bacteriol.* 186, 1430–1437. <https://doi.org/10.1128/JB.186.5.1430-1437.2004>
- Qin, D., Ma, C., Hu, A., Zhang, F., Hu, H., Yu, C.-P., 2016. *Altererythrobacter estronivorus* sp. nov., an estrogen-degrading strain isolated from Yundang lagoon of Xiamen city in China. *Curr. Microbiol.* 72, 634–640. <https://doi.org/10.1007/s00284-016-0995-y>
- Reyes-González, L.V., Hernández de la Cruz, O.N., Castañón-Arreola, M., 2025. Expression of resuscitation-promoting factor C stimulates the growth of *Mycobacterium bovis* BCG and delays DevR regulon activation in hypoxia. *Int. J. Microbiol.* 2025, 2139933. <https://doi.org/10.1155/ijm/2139933>
- Romero, E., Castellanos, J.R.G., Mattevi, A., Fraaije, M.W., 2016. Characterization and crystal structure of a robust cyclohexanone monooxygenase. *Angew. Chem. Int. Ed Engl.* 55, 15852–15855. <https://doi.org/10.1002/anie.201608951>
- Rosłonec, K.Z., Wilbrink, M.H., Capyk, J.K., Mohn, W.W., Ostendorf, M., van der Geize, R., Dijkhuizen, L., Eltis, L.D., 2009. Cytochrome P450 125 (CYP125) catalyses C26-hydroxylation to initiate sterol side-chain degradation in *Rhodococcus jostii* RHA1. *Mol. Microbiol.* 74, 1031–1043. <https://doi.org/10.1111/j.1365-2958.2009.06915.x>

- Sambrook, J., Russell, D.W., 2001. *Molecular Cloning: A Laboratory Manual*. CSHL Press.
- Sang, Y., Xiong, G., Maser, E., 2012. Identification of a new steroid degrading bacterial strain H5 from the Baltic Sea and isolation of two estradiol inducible genes. *J. Steroid Biochem. Mol. Biol.* 129, 22–30. <https://doi.org/10.1016/j.jsbmb.2011.01.018>
- Sangster, J.L., Ali, J.M., Snow, D.D., Kolok, A.S., Bartelt-Hunt, S.L., 2016. Bioavailability and fate of sediment-associated progesterone in aquatic systems. *Environ. Sci. Technol.* 50, 4027–4036. <https://doi.org/10.1021/acs.est.5b06082>
- Schaefer, C.M., Lu, R., Nesbitt, N.M., Schiebel, J., Sampson, N.S., Kisker, C., 2015. FadA5 a thiolase from *Mycobacterium tuberculosis*: a steroid-binding pocket reveals the potential for drug development against tuberculosis. *Struct. Lond. Engl.* 23, 21–33. <https://doi.org/10.1016/j.str.2014.10.010>
- Schiffler, B., Zöllner, A., Bernhardt, R., 2004. Stripping down the mitochondrial cholesterol hydroxylase system, a kinetics study. *J. Biol. Chem.* 279, 34269–34276. <https://doi.org/10.1074/jbc.M402798200>
- Schultes, F.P.J., Welter, L., Hufnagel, D., Heghmanns, M., Kasanmascheff, M., Mügge, C., 2024. An active and versatile electron transport system for cytochrome P450 monooxygenases from the alkane degrading organism *Acinetobacter* sp. OC4. *Chembiochem Eur. J. Chem. Biol.* 25, e202400098. <https://doi.org/10.1002/cbic.202400098>
- Schwindt, A.R., Winkelman, D.L., Keteles, K., Murphy, M., Vajda, A.M., 2014. An environmental oestrogen disrupts fish population dynamics through direct and transgenerational effects on survival and fecundity. *J. Appl. Ecol.* 51, 582–591. <https://doi.org/10.1111/1365-2664.12237>
- Scott, A.P., 2012. Do mollusks use vertebrate sex steroids as reproductive hormones? Part I: Critical appraisal of the evidence for the presence, biosynthesis and uptake of steroids. *Steroids* 77, 1450–1468. <https://doi.org/10.1016/j.steroids.2012.08.009>
- Sharma, S.B., Signer, E.R., 1990. Temporal and spatial regulation of the symbiotic genes of *Rhizobium meliloti* in planta revealed by transposon Tn5-gusA. *Genes Dev.* 4, 344–356. <https://doi.org/10.1101/gad.4.3.344>
- Sim, H.J., Seo, Y.W., Kim, M.R., Song, M.S., Lee, S.Y., 2025. Bisphenol A enhanced cell migration through Kv3.4 in MCF7 cells. *Biochem. Biophys. Res. Commun.* 760, 151678. <https://doi.org/10.1016/j.bbrc.2025.151678>
- Slominski, A.T., Li, W., Kim, T.-K., Semak, I., Wang, J., Zjawiony, J.K., Tuckey, R.C., 2015. Novel activities of CYP11A1 and their potential physiological significance. *J. Steroid Biochem. Mol. Biol.* 151, 25–37. <https://doi.org/10.1016/j.jsbmb.2014.11.010>
- Snapper, S.B., Melton, R.E., Mustafa, S., Kieser, T., Jacobs, W.R., 1990. Isolation and characterization of efficient plasmid transformation mutants of *Mycobacterium smegmatis*. *Mol. Microbiol.* 4, 1911–1919. <https://doi.org/10.1111/j.1365-2958.1990.tb02040.x>
- Souza, B.M., Castro, T.L. de P., Carvalho, R.D. de O., Seyffert, N., Silva, A., Miyoshi, A., Azevedo, V., 2014. σ (ECF) factors of gram-positive bacteria: a focus on *Bacillus subtilis* and the CMNR group. *Virulence* 5, 587–600. <https://doi.org/10.4161/viru.29514>
- Steroids Market size was USD 3451.2 million in 2024 [WWW Document], 2024. . Steroids Mark. Size. URL <https://www.cognitivemarketresearch.com/steroids-market-report> (accessed 4.2.25).
- Stout, E.P., La Clair, J.J., Snell, T.W., Shearer, T.L., Kubanek, J., 2010. Conservation of progesterone hormone function in invertebrate reproduction. *Proc. Natl. Acad. Sci. U. S. A.* 107, 11859–11864. <https://doi.org/10.1073/pnas.1006074107>

- Stover, C.K., de la Cruz, V.F., Fuerst, T.R., Burlein, J.E., Benson, L.A., Bennett, L.T., Bansal, G.P., Young, J.F., Lee, M.H., Hatfull, G.F., 1991. New use of BCG for recombinant vaccines. *Nature* 351, 456–460. <https://doi.org/10.1038/351456a0>
- Strohmaier, S.J., Huang, W., Baek, J.-M., Hunter, D.J.B., Gillam, E.M.J., 2019. Rational evolution of the cofactor-binding site of cytochrome P450 reductase yields variants with increased activity towards specific cytochrome P450 enzymes. *FEBS J.* 286, 4473–4493. <https://doi.org/10.1111/febs.14982>
- Strushkevich, N., MacKenzie, F., Cherkasova, T., Grabovec, I., Usanov, S., Park, H.-W., 2011. Structural basis for pregnenolone biosynthesis by the mitochondrial monooxygenase system. *Proc. Natl. Acad. Sci. U. S. A.* 108, 10139–10143. <https://doi.org/10.1073/pnas.1019441108>
- Swank, R.T., Munkres, K.D., 1971. Molecular weight analysis of oligopeptides by electrophoresis in polyacrylamide gel with sodium dodecyl sulfate. *Anal. Biochem.* 39, 462–477. [https://doi.org/10.1016/0003-2697\(71\)90436-2](https://doi.org/10.1016/0003-2697(71)90436-2)
- Szybalski, W., Kim, S.C., Hasan, N., Podhajski, A.J., 1991. Class-II restriction enzymes--a review. *Gene* 100, 13–26. [https://doi.org/10.1016/0378-1119\(91\)90345-c](https://doi.org/10.1016/0378-1119(91)90345-c)
- Tak, J.D., 1942. On bacteria decomposing cholesterol. *Antonie Van Leeuwenhoek* 8, 32–40. <https://doi.org/10.1007/BF02272764>
- Takyar, S., Hickerson, R.P., Noller, H.F., 2005. mRNA helicase activity of the ribosome. *Cell* 120, 49–58. <https://doi.org/10.1016/j.cell.2004.11.042>
- Tarlera, S., Denner, E.B.M., 2003. *Sterolibacterium denitrificans* gen. nov., sp. nov., a novel cholesterol-oxidizing, denitrifying member of the beta-Proteobacteria. *Int. J. Syst. Evol. Microbiol.* 53, 1085–1091. <https://doi.org/10.1099/ijs.0.02039-0>
- Tarrant, A., Atkinson, M., Atkinson, S., 2004. Effects of steroidal estrogens on coral growth and reproduction. *Mar. Ecol. Prog. Ser.* 269, 121–129. <https://doi.org/10.3354/meps269121>
- Thomas, S.T., Sampson, N.S., 2013. *Mycobacterium tuberculosis* utilizes a unique heterotetrameric structure for dehydrogenation of the cholesterol side chain. *Biochemistry* 52, 2895–2904. <https://doi.org/10.1021/bi4002979>
- Thomas, S.T., VanderVen, B.C., Sherman, D.R., Russell, D.G., Sampson, N.S., 2011. Pathway profiling in *Mycobacterium tuberculosis*: elucidation of cholesterol-derived catabolite and enzymes that catalyze its metabolism. *J. Biol. Chem.* 286, 43668–43678. <https://doi.org/10.1074/jbc.M111.313643>
- Tiso, T., Sabelhaus, P., Behrens, B., Wittgens, A., Rosenau, F., Hayen, H., Blank, L.M., 2016. Creating metabolic demand as an engineering strategy in *Pseudomonas putida* - Rhamnolipid synthesis as an example. *Metab. Eng. Commun.* 3, 234–244. <https://doi.org/10.1016/j.meteno.2016.08.002>
- Tong, W.-Y., Dong, X., 2009. Microbial biotransformation: recent developments on steroid drugs. *Recent Pat. Biotechnol.* 3, 141–153. <https://doi.org/10.2174/187220809788700157>
- Toor, J.S., Sikka, S.C., 2017. Chapter 59 - Developmental and reproductive disorders—role of endocrine disruptors in testicular toxicity, in: Gupta, R.C. *Reproductive and Developmental Toxicology*. Academic Press, pp. 1111–1121. <https://doi.org/10.1016/B978-0-12-804239-7.00059-7>
- Uhía, Galán, B., Medrano, F.J., García, J.L., 2011. Characterization of the KstR-dependent promoter of the gene for the first step of the cholesterol degradative pathway in *Mycobacterium smegmatis*. *Microbiol. Read. Engl.* 157, 2670–2680. <https://doi.org/10.1099/mic.0.049213-0>

- Uhía, I., Galán, B., Morales, V., García, J.L., 2011. Initial step in the catabolism of cholesterol by *Mycobacterium smegmatis* mc2 155. *Environ. Microbiol.* 13, 943–959. <https://doi.org/10.1111/j.1462-2920.2010.02398.x>
- van der Geize, R., Grommen, A.W.F., Hessels, G.I., Jacobs, A. a. C., Dijkhuizen, L., 2011. The steroid catabolic pathway of the intracellular pathogen *Rhodococcus equi* is important for pathogenesis and a target for vaccine development. *PLoS Pathog.* 7, e1002181. <https://doi.org/10.1371/journal.ppat.1002181>
- van Der Geize, R., Hessels, G.I., van Gerwen, R., Vrijbloed, J.W., van Der Meijden, P., Dijkhuizen, L., 2000. Targeted disruption of the *kstD* gene encoding a 3-ketosteroid delta(1)-dehydrogenase isoenzyme of *Rhodococcus erythropolis* strain SQ1. *Appl. Environ. Microbiol.* 66, 2029–2036. <https://doi.org/10.1128/AEM.66.5.2029-2036.2000>
- Vandenberg, L.N., Colborn, T., Hayes, T.B., Heindel, J.J., Jacobs, D.R., Lee, D.-H., Shioda, T., Soto, A.M., vom Saal, F.S., Welshons, W.V., Zoeller, R.T., Myers, J.P., 2012. Hormones and endocrine-disrupting chemicals: low-dose effects and nonmonotonic dose responses. *Endocr. Rev.* 33, 378–455. <https://doi.org/10.1210/er.2011-1050>
- Vasudevan, R., Gale, G.A.R., Schiavon, A.A., Puzorjov, A., Malin, J., Gillespie, M.D., Vavitsas, K., Zulkower, V., Wang, B., Howe, C.J., Lea-Smith, D.J., McCormick, A.J., 2019. CyanoGate: a modular cloning suite for engineering cyanobacteria based on the plant MoClo syntax. *Plant Physiol.* 180, 39–55. <https://doi.org/10.1104/pp.18.01401>
- Waleed, S., Haroon, M., Ullah, N., Tuzen, M., Rind, I.K., Sari, A., 2024. A comprehensive review on advanced trends in treatment technologies for removal of Bisphenol A from aquatic media. *Environ. Monit. Assess.* 197, 83. <https://doi.org/10.1007/s10661-024-13460-x>
- Wang, F.-Q., Yao, K., Wei, D.-Z., 2011. From soybean phytosterols to steroid hormones. *Soybean Health* undefined-undefined. <https://doi.org/10.5772/18808>
- Wang, P., Zheng, D., Liang, R., 2019. Isolation and characterization of an estrogen-degrading *Pseudomonas putida* strain SJTE-1. *3 Biotech* 9, 61. <https://doi.org/10.1007/s13205-018-1537-z>
- Wang, P.-H., Lee, T.-H., Ismail, W., Tsai, C.-Y., Lin, C.-W., Tsai, Y.-W., Chiang, Y.-R., 2013a. An oxygenase-independent cholesterol catabolic pathway operates under oxic conditions. *PloS One* 8, e66675. <https://doi.org/10.1371/journal.pone.0066675>
- Wang, P.-H., Leu, Y.-L., Ismail, W., Tang, S.-L., Tsai, C.-Y., Chen, H.-J., Kao, A.-T., Chiang, Y.-R., 2013b. Anaerobic and aerobic cleavage of the steroid core ring structure by *Steroidobacter denitrificans*. *J. Lipid Res.* 54, 1493–1504. <https://doi.org/10.1194/jlr.M034223>
- Warnke, M., Jacoby, C., Jung, T., Agne, M., Mergelsberg, M., Starke, R., Jehmlich, N., von Bergen, M., Richnow, H.-H., Bröls, T., Boll, M., 2017. A patchwork pathway for oxygenase-independent degradation of side chain containing steroids. *Environ. Microbiol.* 19, 4684–4699. <https://doi.org/10.1111/1462-2920.13933>
- Warnke, M., Jung, T., Jacoby, C., Agne, M., Feller, F.M., Philipp, B., Seiche, W., Breit, B., Boll, M., 2018. Functional characterization of three specific acyl-Coenzyme A synthetases involved in anaerobic cholesterol degradation in *Sterolibacterium denitrificans* Chol1S. *Appl. Environ. Microbiol.* 84, e02721-17. <https://doi.org/10.1128/AEM.02721-17>
- Weber, E., Engler, C., Gruetzner, R., Werner, S., Marillonnet, S., 2011. A modular cloning system for standardized assembly of multigene constructs. *PloS One* 6, e16765. <https://doi.org/10.1371/journal.pone.0016765>
- Weber, S., Leuschner, P., Kämpfer, P., Dott, W., Hollender, J., 2005. Degradation of estradiol and ethinyl estradiol by activated sludge and by a defined mixed culture. *Appl. Microbiol. Biotechnol.* 67, 106–112. <https://doi.org/10.1007/s00253-004-1693-4>

- Wei, W., Fan, S.-Y., Wang, F.-Q., Wei, D.-Z., 2014. Accumulation of androstadiene-dione by overexpression of heterologous 3-ketosteroid Δ^1 -dehydrogenase in *Mycobacterium neoaurum* NwIB-01. *World J. Microbiol. Biotechnol.* 30, 1947–1954. <https://doi.org/10.1007/s11274-014-1614-3>
- WHO Model List of Essential Medicines - 23rd list [WWW Document], 2023. . WHO Model List Essent. Med. - 23rd List. URL <https://www.who.int/publications/i/item/WHO-MHP-HPS-EML-2023.02> (accessed 4.2.25).
- Wijaya, A.J., Dolan, S.K., Kohlstedt, M., Gläser, L., Brear, P., Geddis, S., Wittmann, C., Spring, D.R., Welch, M., 2025. The 2-methylcitrate cycle and the glyoxylate shunt in *Pseudomonas aeruginosa* are linked through enzymatic redundancy. *J. Biol. Chem.* 301, 108355. <https://doi.org/10.1016/j.jbc.2025.108355>
- Wilbrink, M.H., Petrusma, M., Dijkhuizen, L., van der Geize, R., 2011. FadD19 of *Rhodococcus rhodochrous* DSM43269, a steroid-coenzyme A ligase essential for degradation of C-24 branched sterol side chains. *Appl. Environ. Microbiol.* 77, 4455–4464. <https://doi.org/10.1128/AEM.00380-11>
- Winz, C., Zong, W.-X., Suh, N., 2023. Endocrine-disrupting compounds and metabolomic reprogramming in breast cancer. *J. Biochem. Mol. Toxicol.* 37, e23506. <https://doi.org/10.1002/jbt.23506>
- Wipperman, M.F., Yang, M., Thomas, S.T., Sampson, N.S., 2013. Shrinking the FadE proteome of *Mycobacterium tuberculosis*: insights into cholesterol metabolism through identification of an $\alpha\beta\gamma$ heterotetrameric acyl coenzyme A dehydrogenase family. *J. Bacteriol.* 195, 4331–4341. <https://doi.org/10.1128/JB.00502-13>
- Wu, K., Lee, T.-H., Chen, Y.-L., Wang, Y.-S., Wang, P.-H., Yu, C.-P., Chu, K.-H., Chiang, Y.-R., 2019. Metabolites involved in aerobic degradation of the A and B rings of estrogen. *Appl. Environ. Microbiol.* 85, e02223-18. <https://doi.org/10.1128/AEM.02223-18>
- Xu, L.-Q., Liu, Y.-J., Yao, K., Liu, H.-H., Tao, X.-Y., Wang, F.-Q., Wei, D.-Z., 2016. Unraveling and engineering the production of 23,24-bisnorcholesterol steroids in sterol metabolism. *Sci. Rep.* 6, 21928. <https://doi.org/10.1038/srep21928>
- Yam, K.C., D'Angelo, I., Kalscheuer, R., Zhu, H., Wang, J.-X., Snieckus, V., Ly, L.H., Converse, P.J., Jacobs, W.R., Strynadka, N., Eltis, L.D., 2009. Studies of a ring-cleaving dioxygenase illuminate the role of cholesterol metabolism in the pathogenesis of *Mycobacterium tuberculosis*. *PLoS Pathog.* 5, e1000344. <https://doi.org/10.1371/journal.ppat.1000344>
- Yamagishi, M., Sakamoto, T., Ueda, M., Itoh, M., Fujii, Y., Kabumoto, H., Arisawa, A., Fujii, T., 2011. Sterol side chain-cleaving enzyme protein and use thereof. EP2386634A1.
- Yang, B., Zhao, T., Liu, Y., Lu, B., 2023. Guidelines for extraction and quantitative analysis of phytosterols and oxidation products. *eFood* 4, e108. <https://doi.org/10.1002/efd2.108>
- Yang, M., Guja, K.E., Thomas, S.T., Garcia-Diaz, M., Sampson, N.S., 2014. A distinct MaoC-like enoyl-CoA hydratase architecture mediates cholesterol catabolism in *Mycobacterium tuberculosis*. *ACS Chem. Biol.* 9, 2632–2645. <https://doi.org/10.1021/cb500232h>
- Yang, M., Lu, R., Guja, K.E., Wipperman, M.F., St Clair, J.R., Bonds, A.C., Garcia-Diaz, M., Sampson, N.S., 2015. Unraveling cholesterol catabolism in *Mycobacterium tuberculosis*: ChsE4-ChsE5 $\alpha\beta\gamma$ Acyl-CoA dehydrogenase initiates β -oxidation of 3-oxo-cholest-4-en-26-oyl CoA. *ACS Infect. Dis.* 1, 110–125. <https://doi.org/10.1021/id500033m>
- Yang, W., Bell, S.G., Wang, H., Zhou, W., Hoskins, N., Dale, A., Bartlam, M., Wong, L.-L., Rao, Z., 2010. Molecular characterization of a Class I P450 electron transfer system from *Novosphingobium aromaticivorans* DSM12444. *J. Biol. Chem.* 285, 27372–27384. <https://doi.org/10.1074/jbc.M110.118349>

- Yang, X., Dubnau, E., Smith, I., Sampson, N.S., 2007. Rv1106c from *Mycobacterium tuberculosis* is a 3 β -hydroxysteroid dehydrogenase. *Biochemistry* 46, 9058–9067. <https://doi.org/10.1021/bi700688x>
- Yin, G.-G., Kookana, R.S., Ru, Y.-J., 2002. Occurrence and fate of hormone steroids in the environment. *Environ. Int.* 28, 545–551. [https://doi.org/10.1016/s0160-4120\(02\)00075-2](https://doi.org/10.1016/s0160-4120(02)00075-2)
- Yoshimoto, T., Nagai, F., Fujimoto, J., Watanabe, K., Mizukoshi, H., Makino, T., Kimura, K., Saino, H., Sawada, H., Omura, H., 2004. Degradation of estrogens by *Rhodococcus zopfii* and *Rhodococcus equi* isolates from activated sludge in wastewater treatment plants. *Appl. Environ. Microbiol.* 70, 5283–5289. <https://doi.org/10.1128/AEM.70.9.5283-5289.2004>
- Yu, C.-P., Roh, H., Chu, K.-H., 2007. 17 β -estradiol-degrading bacteria isolated from activated sludge. *Environ. Sci. Technol.* 41, 486–492. <https://doi.org/10.1021/es060923f>
- Zhang, R., Zhang, Y., Wang, Y., Yao, M., Zhang, J., Liu, H., Zhou, X., Xiao, W., Yuan, Y., 2019. Pregnenolone overproduction in *Yarrowia lipolytica* by integrative components pairing of the cytochrome P450_{scc} system. *ACS Synth. Biol.* 8, 2666–2678. <https://doi.org/10.1021/acssynbio.9b00018>
- Zhang, T., Xiong, G., Maser, E., 2011. Characterization of the steroid degrading bacterium S19-1 from the Baltic Sea at Kiel, Germany. *Chem. Biol. Interact.* 191, 83–88. <https://doi.org/10.1016/j.cbi.2010.12.021>
- Zhao, A., Zhang, X., Li, Y., Wang, Z., Lv, Y., Liu, J., Alam, M.A., Xiong, W., Xu, J., 2021. *Mycolicibacterium* cell factory for the production of steroid-based drug intermediates. *Biotechnol. Adv.* 53, 107860. <https://doi.org/10.1016/j.biotechadv.2021.107860>
- Zhao, Y., Ye, L., Zhang, X.-X., 2018. Emerging Pollutants—Part I: occurrence, fate and transport. *Water Environ. Res.* 90, 1301–1322. <https://doi.org/10.2175/106143018X15289915807236>
- Zhao, Y.-Q., Liu, Y.-J., Song, L., Yu, D., Liu, Kun, Liu, Ke, Gao, B., Tao, X.-Y., Xiong, L.-B., Wang, F.-Q., Wei, D.-Z., 2023. Unravelling and engineering an operon involved in the side-chain degradation of sterols in *Mycolicibacterium neoaurum* for the production of steroid synthons. *Biotechnol. Biofuels Bioprod.* 16, 121. <https://doi.org/10.1186/s13068-023-02376-2>
- Zhou, Y., Zha, J., Xu, Y., Lei, B., Wang, Z., 2012. Occurrences of six steroid estrogens from different effluents in Beijing, China. *Environ. Monit. Assess.* 184, 1719–1729. <https://doi.org/10.1007/s10661-011-2073-z>
- Zhou, Z., Guo, Q., Xu, Z., Wang, L., Cui, K., 2015. Distribution and removal of endocrine-disrupting chemicals in industrial wastewater treatment. *Environ. Eng. Sci.* 32, 203–211. <https://doi.org/10.1089/ees.2014.0082>



X. ANEXOS

ANEXO I. Publicación 1

RESEARCH ARTICLE

Identification of the aldolase responsible for the production of 22-hydroxy-23,24-bisnorchol-4-ene-3-one from natural sterols in *Mycolicibacterium smegmatis*

Gabriel Hernández-Fernández | Miguel G. Acedos | José L. García  | Beatriz Galán 

Department of Microbial and Plant Biotechnology, Centro de Investigaciones Biológicas Margarita Salas (CSIC), Madrid, Spain

Correspondence

Beatriz Galán, Department of Microbial and Plant Biotechnology, Centro de Investigaciones Biológicas Margarita Salas (CSIC), Madrid, Spain.
Email: bgalan@cib.csic.es

Funding information

Comunidad de Madrid, Grant/Award Number: S2018/BAA-4532; Grant, Grant/Award Number: FPU21/05101; Ministerio de Ciencia e Innovación, Grant/Award Number: PID2019-110612RB-I00 and PID2021-125370OB-I00; Ministerio de Economía y Competitividad

Abstract

Mycobacterial mutants blocked in ring degradation constructed to achieve C19 synthons production, also accumulate by-products such as C22 intermediates throughout an alternative pathway reducing the production yields and complicating the downstream purification processing of final products. In this work, we have identified the *MSMEG_6561* gene, encoding an aldolase responsible for the transformation of 22-hydroxy-3-oxo-cholest-4-ene-24-carboxyl-CoA (22-OH-BCN-CoA) into the 22-hydroxy-23,24-bisnorchol-4-ene-3-one (4-HBC) precursor (20S)-3-oxopregn-4-ene-20-carboxaldehyde (3-OPA). The deletion of this gene increases the production yield of the C-19 steroidal synthon 4-androstene-3,17-dione (AD) from natural sterols, avoiding the production of 4-HBC as by-product and the drawbacks in the AD purification. The molar yield of AD production using the MS6039-5941-6561 triple mutant strain was checked in flasks and bioreactor improving very significantly compared with the previously described MS6039-5941 strain.

INTRODUCTION

One of the most attractive and efficient strategies for producing steroid precursors useful for the synthesis of steroid pharmaceutical drugs is to utilize mycobacterial strains to transform natural sterols into useful intermediates or synthons as building blocks (Donova, 2021; Fernández-Cabezón et al., 2018; Nunes et al., 2022). Currently, three main C19 synthons are produced at industrial scale by different mycobacterial strains, i.e., 4-androstene-3,17-dione (AD), 1,4-androstadiene-3,17-dione (ADD) and 9 α -hydroxy-4-androstenedione (9OH-AD) using phytosterols as natural precursors. C19 synthons are used in the chemical synthesis of sex hormones such as testosterone, estrone, oestradiol and mineralocorticoids (Feng et al., 2022; Zhao et al., 2021). Apart from traditional

mutation approaches, metabolic engineering techniques have been developed recently to manipulate the metabolism of sterols in these bacteria in order to accumulate the intermediate synthons at high yields in the culture medium (Fernández-Cabezón et al., 2018; Wang et al., 2022).

The aerobic sterol degradation pathway of mycobacterial strains has been studied extensively and most of the important key intermediates have been identified. It begins with the transport of sterols to the cellular interior mediated by the system, encoded by the *mce* operon (Pandey & Sassetti, 2008). Sterol degradation involves two main processes: the removal of the side chain and the cleavage of steroid rings. After the sterols have entered into the cell, two sequential reactions take place: the oxidation of cholesterol to 5-cholesten-3-one and its isomerization to 4-cholesten-3-one catalysed

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs](https://creativecommons.org/licenses/by-nc-nd/4.0/) License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2023 The Authors. *Microbial Biotechnology* published by Applied Microbiology International and John Wiley & Sons Ltd.

by a 3β -hydroxysteroid dehydrogenase/isomerase (3β -HSD) that uses NAD or NADP as cofactors (Uhía et al., 2011). The aliphatic chain of sterols is degraded simultaneously, following a mechanism similar to β -oxidation of fatty acids, by three β -cyclic oxidations mediated by cytochromes P450 CYP125 and CYP142 (Capyk et al., 2009; Ouellet et al., 2010). Each cycle comprises enzymes such as ligases, dehydrogenases, hydratases and one aldolase complex responsible of the AD formation and the release of two propionyl-CoA and one acetyl-CoA molecules (Figure 1) (García et al., 2012). Considering this information, it has been demonstrated that the knockout of the *kstD*, *kshA* and *kshB* genes in different combinations allows the production of AD, ADD and 9OH-AD (Fernández-Cabezón et al., 2018; Wang et al., 2022).

Although well studied, there are still unknown aspects of *Mycobacterium* metabolism and in particular of *M. smegmatis* that need to be elucidated in order to achieve efficient industrial production of some intermediaries. For instance, the knockout of the *kstD* (*MSMEG_5941*) and *kshB* (*MSMEG_6039*) yielded an AD producer strain but it accumulates a certain amount of the by-product 4-HBC, which reduces the yield and

hinders the purification of the final product (Galán et al., 2017).

Considering that *M. smegmatis* does not have an active *mnOpccR* reductase pathway that could explain the production of 4-HBC (Peng et al., 2021) (see below) (Figure 1), we propose that most probably this compound is produced in *M. smegmatis* by the aldolytic reaction to transform of 22-hydroxy-3-oxo-cholest-4-ene-24-carboxyl-CoA (22-OH-BCN-CoA) into (20S)-3-oxopregn-4-ene-20-carboxaldehyde (3-OPA) that is the proposed precursor of 4-HBC (Figure 1). However, the enzymes involved in this transformation in *M. smegmatis* remained unknown. Retroaldol elimination of acetyl-CoA has been described in the degradation of the cholic acid side/chain in *Pseudomonas* sp. Chol1, renamed as *Stutzerimonas degradans* (Gomila et al., 2022; Holert et al., 2013). Both compounds, 22-OH-BCN-CoA, precursor of 4-HBC, and cholic acid are structurally similar and this reaction is catalysed by Sal (steroid aldolase), a retro-aldolase (Holert et al., 2013).

In this work, we have identified the *MSMEG_6561* gene encoding an aldolase enzyme responsible for the transformation of 22-OH-BCN-CoA into 3-OPA

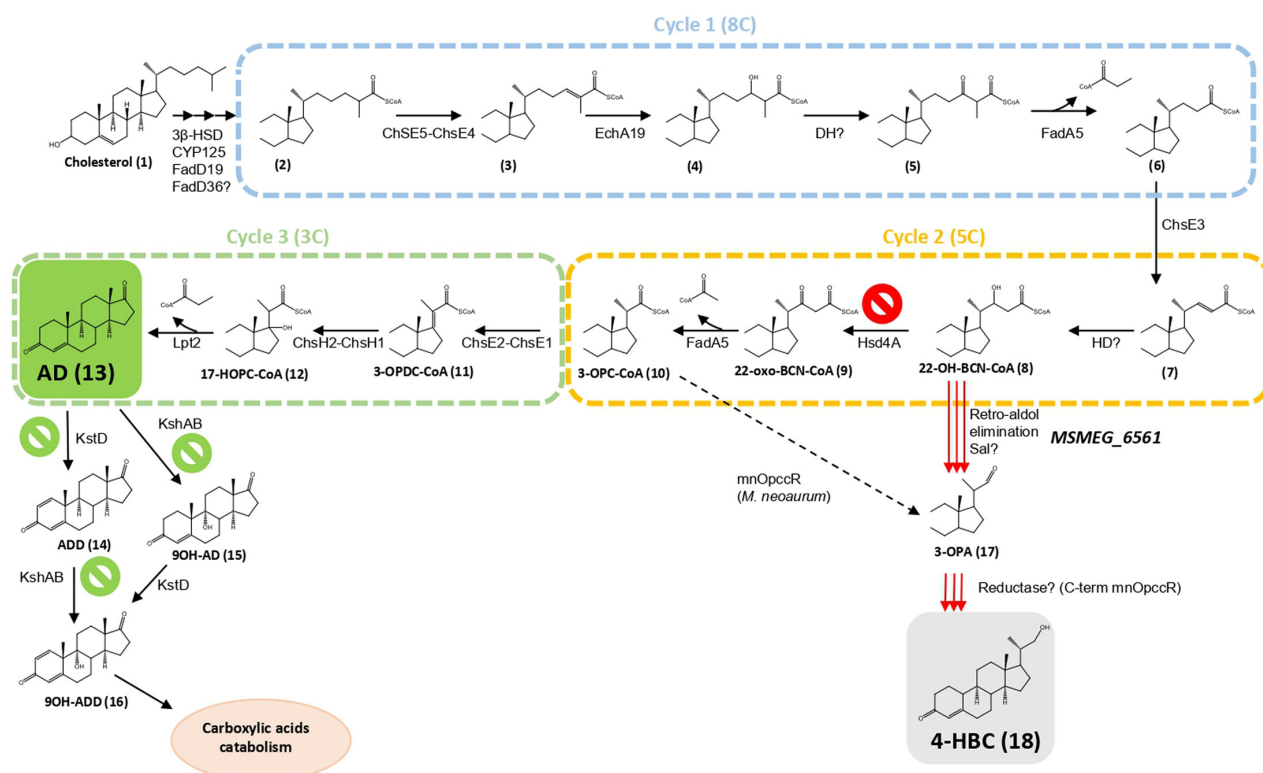


FIGURE 1 Cholesterol side-chain degradation pathway in Actinobacteria. Green symbols indicate the mutations performed to accumulate the C19 compounds. The three β -oxidation-like cycles are indicated with boxes. Enzymes with question marks are unknown and the biochemical steps have not been verified. DH, dehydrogenase; HD, hydratase; Sal, steroid aldolase. (1) Cholesterol, (2) (25S)-cholest-4-en-3-one-26-oil-CoA, (3) Cholest-4,24-dien-3-one-26-oil-CoA, (4) cholest-4-en-24-ol-3-one-26-oil-CoA, (5) 3,24-dioxocholest-4-en-26-oil-CoA, (6) 3-oxochole-4-en-24-oil-CoA, (7) 3-oxochole-4,22-dien-24-oil-CoA, (8) 22-hydroxyl-3-oxochole-4-en-24-oil-CoA, (9) 3,22-dioxochole-4-en-24-oil-CoA, (10) 3-oxo-23,24-bischole-4-en-22-oil-CoA, (11) 3-oxo-4,17-pregnadiene-20-carboxyl-CoA, (12) 3-oxo-23,24-bischole-4-en-17-ol-22-oil-CoA, (13) 4-androstene-3,17-dione, (14) 1,4-androstadiene-3,17-dione, (15) 9 α -hydroxy-4-androstenedione, (16) 9OHADD, (17) 3-OPA, (18) 4-HBC.

in *M. smegmatis*. We have demonstrated that a triple MS6039-5941-6561 mutant of *M. smegmatis* is unable to produce 4-HBC and therefore, it produces AD from natural sterols with a significant higher yield than the previous AD producer mutant.

EXPERIMENTAL PROCEDURES

Chemicals

4-Androstene-3,17-dione was purchased from TCI America. Chloroform and glycerol were purchased from Merck. Cholesterol, cholestenone, N,O-bis(trimethylsilyl) trifluoroacetamide (BSTFA), gentamicin, pyridine, Tween 80 and tyloxapol were from Sigma. Oligonucleotides were from Sigma-Genosys. Phytosterols were provided by Gadea Biopharma containing a mixture of different sterols (w/w percentage): brassicasterol (2.16%), stigmasterol (8.7%), campesterol (36.8%) and β -sitosterol (54.4%).

Bacterial strains and culture conditions

The strains as well as the plasmids and primers used in this work are listed in Table 1. Preinocula of *M. smegmatis* mc² 155 and its derivative strains were grown in LB medium with 0.05% Tween 80 (Sigma) at 37°C with orbital shaking at 200 rpm. LB agar plates and 7H10 agar (Difco) with sucrose 10% (w/v) were used for solid media. *Escherichia coli* strains were grown in LB medium at 37°C with orbital shaking at 200 rpm and in LB agar plates for solid media.

For the biotransformation assays in 100 mL flasks, Middlebrook 7H9 broth medium (Difco) supplemented with 9 mM glycerol (Sigma) as carbon and energy source and 0.4 g/L cholesterol (Sigma) as raw material were used. Due to the high hydrophobicity of cholesterol, the 5 mM stock was prepared in tyloxapol 10% (v/v) (Sigma), sonicated for 30 min and autoclaved twice until its total dissolution. Tyloxapol final concentration in the biotransformation medium was 2%.

The preinoculum was first grown over-weekend on LB with Tween 80 as described before. Then, it was washed twice with saline solution 0.85% with 0.05% Tween 80 and the biotransformation assays were inoculated at $OD_{600}=0.05$ in 25 mL of biotransformation media. Strains were grown at 200 rpm and 37°C for 96 h and 500 μ L samples of the culture were taken every 24 h and kept at -20°C until their analysis by GC/MS.

Bioreactor experiments were performed using the culture media previously described (Kutney et al., 2003) containing: corn steep liquor (20 g/L), sodium nitrate (5.4 g/L), ammonium chloride (0.28 g/L) and monobasic sodium phosphate (0.28 g/L). 10 g of phytosterols were

dissolved in 160 mL of sunflower oil and mixed with the aqueous phase (740 mL). The inoculum volume was 10%. This experiment was carried out in a commercial stirred tank bioreactor (STBR) of 2 L total volume (Sartorius AG) and a working volume of 1 L was used. The process was carried out at 37°C, pH was adjusted at the start of the reaction to 7.4 and was not controlled during the process, 600 rpm of stirrer speed and an air flow rate of 1 vvm were used. Dissolved oxygen was maintained above 20% throughout the process. When needed, final concentrations of 5 μ g/mL of gentamicin and 20 μ g/mL of kanamycin were added to the growth media for *M. smegmatis* strains, while 10 μ g/mL of gentamicin and 50 μ g/ μ L of kanamycin were used for the *E. coli* strains.

Gene deletion

The knock-out strain of *M. smegmatis* named MS6039-5941-6561 was constructed by homologous recombination using the pJQ200x plasmid, a derivative of the suicide pJQ200 vector that does not replicate in *Mycobacterium* (Jackson et al., 2001). The strategy consists in generating two fragments of ~600 bp, the first one containing the upstream region and few nucleotides of the 5'-end of the gene and the second one containing the downstream region and few nucleotides of its 3'-end of the gene, that are amplified by PCR using the oligonucleotides described in Table 1 and *M. smegmatis* genomic DNA as template. *M. smegmatis* genomic DNA extraction was performed as described (Uhía et al., 2011). The two fragments generated were digested with the corresponding enzymes and cloned into the plasmid pJQ200x using *E. coli* DH10B competent cells as described (Uhía et al., 2011). Plasmid DNA from *E. coli* DH10B recombinant strain was extracted using the High Pure Plasmid Purification Kit (Roche), according to the manufacturer's instructions. This procedure was performed for gene *MSMEG_6561* generating the pJQ6561 plasmid. Plasmid pJQ6561 was electroporated into competent *M. smegmatis* MS6039-5941 to obtain the strain MS6039-5941-6561. Single cross-overs were selected on 7H10 agar plates containing gentamicin and the presence of the *xylE* gene encoded in pJQ200x was confirmed by spreading catechol over the single colonies of electroporated *M. smegmatis*. The appearance of a yellow coloration indicates the presence of the *xylE* gene. Colonies were also contra-selected in 10% sucrose. A single colony was grown in 10 mL of LB medium with 5 μ g/mL gentamicin up to an optical density of 0.8–0.9 and 20 μ L of a 1:100 dilution was plated onto 7H10 agar plates with 10% sucrose to select for double cross-overs. Potential double cross-overs (sucrose-resistant colonies) were screened for gentamicin

sensitivity and the absence of the *xyIE* gene. The mutant strain MS6039-5941-6561 was analysed by PCR and DNA sequencing to confirm the deletions of *MSMEG_6561* gene.

Construction of pMV6561 plasmid

MSMEG_6561 gene from *M. smegmatis* mc² 155 genomic DNA was chemically synthesized by GeneScript. The synthetic gene was digested with *Bam*HI and *Hind*III and cloned into pMV261, a shuttle plasmid that replicates in *E. coli* and *M. smegmatis* generating the plasmid pMV6561 that was transformed into *E. coli* DH10B to generate the recombinant strain *E. coli* DH10B (pMV6561). Once the sequence of the pMV6561 was checked it was used to transform electro-competent cells of *M. smegmatis* MS6039-5941-6561 generating the *M. smegmatis* MS6039-5941-6561 (pMV6561) recombinant strain.

GC/MS analyses

To perform GC/MS analysis, culture aliquots (0.2 mL) were extracted twice at various extents of incubation with an equal volume of chloroform. Previously to its extraction, 10 µL of a solution of 10 mM cholesterol (when phytosterols were used as substrate) or 10 mM cholestenone (when cholesterol was used as substrate) dissolved in chloroform were added to the aliquots as internal standards. The chloroform fraction was concentrated by evaporation and the trimethylsilyl ether derivatives were formed by reaction with 50 µL of TMT-BSTFA and 50 µL of pyridine and heating at 80°C for 20 min. Calibration standards were derivatized in the same way. The GC/MS analysis was carried out using an Agilent 7890A gas chromatograph coupled to an Agilent 5975C mass detector (Agilent Technologies). Mass spectra were recorded in electron impact (EI) mode at 70 eV within the *m/z* range 50–550. The chromatograph was equipped with a 30 m × 0.25 mm i.d. capillary column (0.25 µm film thickness) HP-5MS (5% diphenyl 95% dimethylpolysiloxane from Agilent Technologies). Working conditions in the sample were as follows: split ratio (20:1), injector temperature, 320°C; column temperature 240°C for 3 min, then heated to 320°C at 5°C/min. For quantification of the peak area, the quantitative masses were 329 + 458 *m/z* for cholesterol; 343 + 384 *m/z* for cholestenone; 382 + 472 *m/z* for campesterol; 394 + 484 *m/z* for stigmasterol; 357 + 486 *m/z* for β-sitosterol; 244 + 286 *m/z* for AD and 73 + 270 *m/z* for 4-HBC in selected ions of monitoring. The mass spectra and retention data were used to assess the identity of compounds by comparing them with those of standards in the NIST Mass

Spectral Database and commercial standards (NIST 2011).

Calculations

The molar yield (MY) of the products were determined according to the following equation (Equation 1):

$$MY (\%) = \frac{[AD](mM)}{[Sc](mM)} \cdot 100 \quad (1)$$

In the equation, the term “AD” refers to 4-androstene-3,17-dione and “Sc” refers to consumed sterols.

The productivity (*P*) of 4-androstene-3,17-dione (AD) was determined according to the following equation (Equation 2):

$$P \left(\frac{g}{L \cdot h} \right) = \frac{[AD](g/L)}{t(h)} \quad (2)$$

In the equation, the term “*t*” refers to time (h).

RESULTS AND DISCUSSION

Identification of the aldolase-encoding gene in the genome of *M. smegmatis*

Mutants constructed to obtain C19 synthons accumulate C22 intermediate by-products, reducing yield and making purification of the final product difficult (Galán et al., 2017). In the same way, mutants with the blocked side chain degradation pathway constructed to produce C22 steroids show accumulation of C19 residues (Galán et al., 2017; Xu et al., 2016). Thus, looking to the side-chain sterol degradation pathway we can find a conventional route by which AD and other C19 steroids are generated as degradation intermediates, and an alternative route that gives rise to 4-HBC and other C22 steroids as final degradation products (Song et al., 2023). The central node of both pathways is 22-OH-BCN-CoA, which leads to AD or to 4-HBC through two possible pathways, one proposed in *Mycobacterium neoaurum* (Peng et al., 2021) and an unknown aldolytic reaction (Figure 1).

The pathway described in *M. neoaurum* starts from 3-oxo-4-pregnen-20-carboxyl-CoA (22-OPC-CoA) that is generated from 22-OH-BCN-CoA by the consecutive action of two enzymes, i.e., Hsd4A (a dual-function enzyme, with both 17β-hydroxysteroid dehydrogenase and β-hydroxyacyl-CoA dehydrogenase activities) (Xu et al., 2016) and FadA5 thiolase (Nesbitt et al., 2010) (Figure 1). Next the transformation of 22-OPC-CoA into 4-HBC is made by means of

a dual function reductase named mnOpccR. mnOpccR acts catalysing both a 4e⁻ reduction reaction and a 2e⁻ reduction reaction for the formation of 4-HBC. For the 4e⁻ reduction, a first 2e⁻ reduction takes place at the C-terminus of the enzyme, to convert the 3-OPC-CoA to 3-oxo-4-pregnen-20-carboxyl aldehyde (3-OPA) followed by a second 2e⁻ reduction, at the N-terminus, of the aldehyde group to alcohol to form 4-HBC. In the 2e⁻ reduction reaction, the only substrate is 3-OPA, which is reduced at the N-terminus of mnOpccR. It has been shown that inactivation or overexpression of mnOpccR in *M. neoaurum* can achieve exclusive production of AD or 4-HBC (Peng et al., 2021).

In *M. smegmatis*, a gene homologous to *mnopccR* was found with 78% identity, i.e., *MSMEG_1623* (accession number: AIU06854). However, Peng et al. (2021) have demonstrated that the C-terminal end of the *MSMEG_1623* protein is not active. Therefore, the only way to obtain 4-HBC in *M. smegmatis* would be through the degradation of the 22-OH-BCN-CoA through an alternative route. This conversion could occur via a retroaldol elimination, or aldolytic cleavage, that would produce 3-OPA (Figure 1), followed by a reduction by the N-terminal domain of msOpccR reductase to generate 4-HBC (Peng et al., 2021) or by another reductase activity.

The retroaldol elimination of acetyl-CoA has been described in the degradation of the cholic acid side chain in *S. degradans* (Holert et al., 2013). Both compounds, 22-OH-BCN-CoA, precursor of 4-HBC, and cholic acid are structurally similar. This reaction is catalysed by the enzyme called Sal (steroid aldolase), encoded by the *sal* gene (Holert et al., 2013). Chemically, the aldolytic reaction of 22-OH-BCN-CoA (24C) to 3-OPA (22C aldehyde) could be performed by a *sal*-encoded aldol-lyase similar to that of the degradation of cholate side chain in *S. degradans*.

To identify the gene responsible for encoding this activity in *M. smegmatis*, we searched using Blastp those sequences similar to the Sal enzyme within its genome. Thus, the *MSMEG_6561* gene was identified, whose amino acid sequence shares a 57% similarity with the Sal enzyme (OEE10244.1) of *S. degradans*. Other *Mycolicibacterium* such as *M. neoaurum* or *M. fortuitum* that has been used for the production of steroid synthons also encodes a homologous Sal enzyme (Table 2). Moreover, *Mycobacterium tuberculosis* has also a homologous protein (Table 2).

Remarkably, the *MSMEG_6561* gene is expressed constitutively in *M. smegmatis* (Uhía et al., 2012). The transcriptomic analysis carried out in the presence or absence of cholesterol indicates that *MSMEG_6561* is not induced in the presence of cholesterol suggesting that it could play other role in the metabolism of some related compounds.

MSMEG_6561 gene is part of a putative operon containing *MSMEG_6560*, encoding a putative

enoyl-CoA-hydratase, *MSMEG_6562*, encoding a putative acyl-CoA dehydrogenase, and *MSMEG_6563* coding a putative oxidoreductase having Rossmann-fold NAD(P)-binding domain-containing protein according to the Conserved Domain DB at NCBI. The resemblance of this operon to that of the *S. degradans* strain, where the aldolase Sal and hydratase Shy coding genes are located contiguous in the genome (Holert et al., 2013), suggests that it could very likely to be involved in the formation of 3-OPA and subsequent reduction to 4-HBC in the MS6039-5941 mutant strain. Elucidation of the specific role of this operon in the metabolism of *M. smegmatis* requires further studies but most probably could be involved in the degradation of some related compounds such as bile acids and/or fatty acids.

Construction of the triple MS6039-5941-6561 mutant

To study the possible aldolase activity of the *MSMEG_6561* gene, we have mutated the strain *M. smegmatis* MS6039-5941 ($\Delta kshB1$, $\Delta kstD1$) which is the AD producer that accumulates 4-HBC as a by-product (Galán et al., 2017) by double homologous recombination yielding the triple mutant *M. smegmatis* MS6039-5941-6561 ($\Delta kshB1$, $\Delta kstD1$, Δsal). The mutation of the *MSMEG_6561* gene was confirmed by PCR and DNA sequencing by Sanger.

The MS6039-5941-6561 mutant is able to grow in minimal medium using glycerol or glucose as carbon and energy sources as efficiently as the MS6039-5941 mutant, indicating that the mutation of *MSMEG_6561* gene does not affect the growth in these substrates (data not shown). In addition, the morphology of the mutant cells and the plate colonies does not appear to be affected.

Production of AD in the *M. smegmatis* MS6039-5941-6561 strain

The mutant strains MS6039-5941 and MS6039-5941-6561 were grown in minimal medium with glycerol as energy and carbon sources and in the presence of 0.4g/L cholesterol to address the production of steroid intermediates in flasks. Figure S1 shows that the triple mutation has not impaired strain growth in the biotransformation medium compared with the MS6039-5941 strain. Steroids were extracted from aliquots of the culture medium every 24 h for 96 h and were separated and identified by GC/MS, as detailed in Materials and Methods. As previously mentioned, the degradation of the sterol side chain can follow two possible routes in *M. smegmatis*: the conventional one by which AD and other C-19 steroids are generated

as degradation intermediates and an alternative route that gives rise to the 4-HBC and other C-22 steroids as degradation end products (Figure 1). Figure 2 shows the chromatograms and mass spectra obtained from the samples of MS6039-5941 and MS6039-5941-6561 strain, respectively, taken at 96 h. While in the case of MS6039-5941 strain, three major peaks are observed: (1 and 2) corresponding to the accumulated products: AD and 4-HBC, respectively, and (4) corresponding to 4-cholesten-3-one (internal standard, hereinafter cholestenone). There is not 4-HBC detected when the triple mutant strain was used. Other minor peaks that have been identified as non-steroidal trimethylsilyl residues of the reagents used in the derivatization of the samples are observed. MS6039-5941-6561 strain reached

up to 240 mg/L of AD after 96 h, while MS6039-5941 only produced up to 150 mg/L (Figure 3A). Molar yields of the biotransformation process after 96 h reached 59% and 81% for the MS6039-5941 and MS6039-5941-6561 strains, respectively (Figure 3B).

To further confirm that *MSMEG_6561* is the coding gene responsible of the accumulation of 4-HBC during the cholesterol biotransformation, the MS6039-5941-6561 strain was transformed with plasmid pMV6561 carrying the *MSMEG_6561* gene under the control of P_{hsp60} promoter yielding the MS6039-5941-6561 (pMV6561) strain (Table 1). The resulting strain was checked in a biotransformation assay in the presence of 0.4 g/L of cholesterol. First, the ability to grow in the biotransformation medium at the same

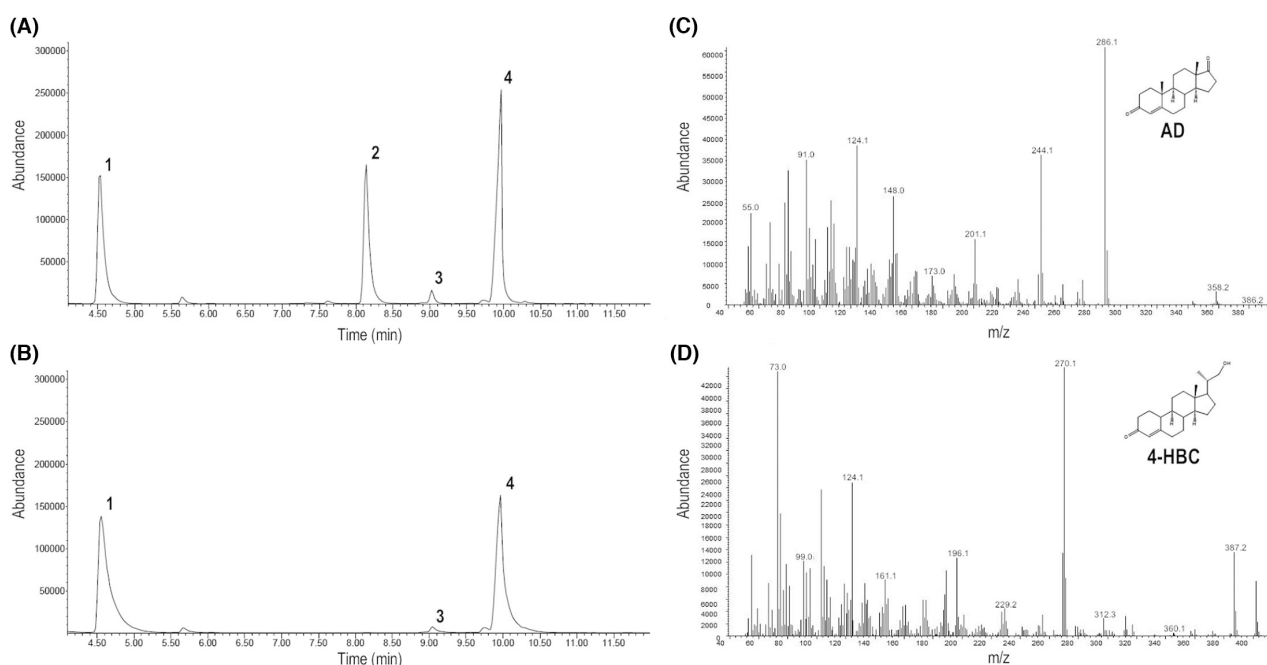


FIGURE 2 (A) Analysis by GC/MS of the products accumulated in the culture of MS6039-5941 and (B) MS6039-5941-6561 strains after 96 h in a medium containing 9 mM glycerol and cholesterol at 0.4 g/L. (1) 4-Androstene-3,17-dione (AD), (2) 4-HBC, (3) cholesterol and (4) cholestenone (internal standard). (C) Mass spectra and fragmentation pattern of AD and (D) 4-HBC.

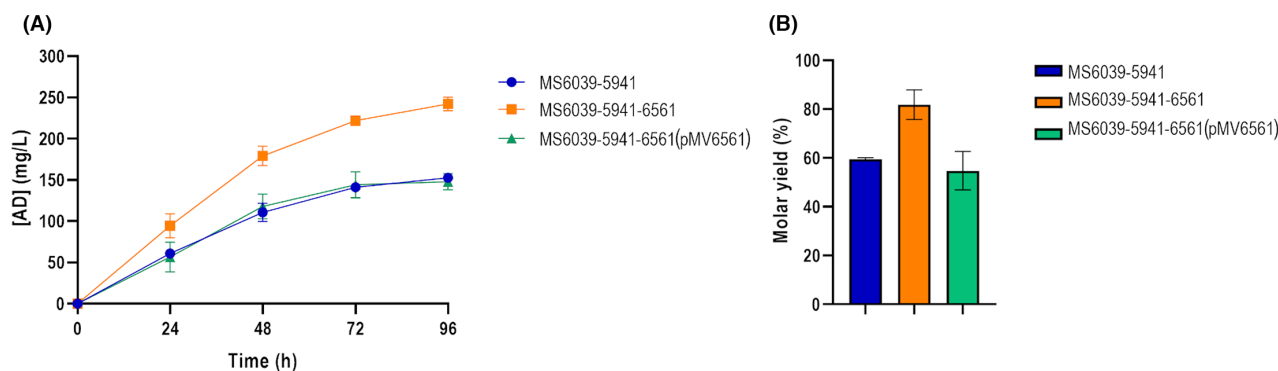


FIGURE 3 (A) 4-Androstene-3,17-dione (AD) production of the MS6039-5941, MS6039-5941-6561 and MS6039-5941-6561 (pMV6561) strains cultured on shake flasks in a medium containing 9 mM glycerol and cholesterol at 0.4 g/L. (B) Molar yield of AD production of MS6039-5941, MS6039-5941-6561 and MS6039-5941-6561 (pMV6561) strains.

TABLE 1 Strains, plasmids and primers.

Genotype and/or description		Reference
Strain		
<i>Escherichia coli</i> DH10B	F ⁻ , <i>mcrA</i> , Δ (<i>mrrhsdRMS-mcrBC</i>), Φ80dlacZDM15, Δ <i>lacX74</i> , <i>deoR</i> , <i>recA1</i> , <i>araD139</i> , Δ(<i>ara-leu</i>)7697, <i>galU</i> , <i>galK</i> , λ ⁻ , <i>rpsL</i> , <i>endA1</i> , <i>nupG</i>	Invitrogen
<i>Mycobacterium smegmatis</i> mc ² 155	<i>ept-1</i> , mc ² 6 mutant efficient for electroporation	Snapper et al. (1990)
MS6039-5941	<i>M. smegmatis</i> mc ² 155 Δ <i>MSMEG_6039</i> (Δ <i>kshB1</i>) and Δ <i>MSMEG_5941</i> (Δ <i>kstD1</i>)	Galán et al. (2017)
MS6039-5941-6561	<i>M. smegmatis</i> mc ² 155 Δ <i>MSMEG_6039</i> (Δ <i>kshB1</i>), Δ <i>MSMEG_5941</i> (Δ <i>kstD1</i>) and Δ <i>MSMEG_6561</i> (Δ <i>Sal</i>)	This work
MS6039-5941-6561 (pMV6561)	<i>M. smegmatis</i> mc ² 155 Δ <i>MSMEG_6039</i> (Δ <i>kshB1</i>), Δ <i>MSMEG_5941</i> (Δ <i>kstD1</i>) and Δ <i>MSMEG_6561</i> (Δ <i>Sal</i>) with pMV6561	This work
MS6039-5941-6561 (pMV261)	<i>M. smegmatis</i> mc ² 155 Δ <i>MSMEG_6039</i> (Δ <i>kshB1</i>), Δ <i>MSMEG_5941</i> (Δ <i>kstD1</i>) and Δ <i>MSMEG_6561</i> (Δ <i>Sal</i>) with pMV261	This work
Plasmids		
pJQ200x	Suicide vector used to perform directed mutagenesis in <i>Mycobacterium</i> , P15A ori, <i>sacB</i> , Gm ^r , <i>xylE</i>	Jackson et al. (2001)
pJQ6561	pJQ200x containing one fragment upstream and another downstream of <i>MSMEG_6561</i> between XbaI-SacI	This work
pMV261	<i>Mycobacterium/E. coli</i> shuttle vector with the kanamycin resistance <i>aph</i> and the promoter from the <i>hsp60</i> gene from <i>M. tuberculosis</i>	Stover et al. (1991)
pMV6561	pMV261 harbouring the <i>MSMEG_6561</i> synthetic gene encoding <i>Sal</i> from <i>M. smegmatis</i> mc ² 155 between BamHI-HindIII	This work
Primers		
MSMEG_6561_UP_F MSMEG_6561_UP_R	CGGGGAGCTCGTCGGGACGGTCGCAGAACACG CCCTGGATCCACGACCTGACCTGACCTCCGCGGC	Amplification of fragment upstream MSMEG_6561 (SacI-BamHI)
MSMEG_6561_DOWN_F MSMEG_6561_DOWN_R	ACTTGGATCCATTTGCCCGCACGCCGAGGT GGCTTCTAGAACATCGACCGCATGGGTGCCCT	Amplification of fragment downstream MSMEG_6561 (BamHI-XbaI)

TABLE 2 Proteins homologous to MSMEG_6561.

Accession N°	Product	Identity (%)	Coverage (%)	E-value	Strain
ADM35101.1	Steroid ketothiolase	57.1	98	e-160	<i>Stutzerimonas degradans</i>
WP_171511544.1	Thiolase family protein	89.9	100	0	<i>Mycobacterium fortuitum</i>
WP_023985643.1	Thiolase family protein	71.0	98	0	<i>Mycobacterium neoaurum</i>
TXA42495.1	Thiolase family protein	95.0	98	0	<i>Mycobacterium tuberculosis</i>

rate than the parental strain was confirmed (Figure S1). The GC–MS analysis of the supernatant shows that the complementation with *MSMEG_6561* gene allows the formation of the by-product 4-HBC (Figure S2). MS6039-5941-6561 (pMV6561) strain showed the same levels of AD production as MS6039-5941 (around 150 mg/L) and the molar yield after 96 h reached 60% (Figure 3).

To determine the behaviour of the new strain using phytosterols as a substrate, a biotransformation assay was carried out in a STBR. The MS6039-5941-6561 mutant was grown for 96 h and samples were taken every 24 h. Figure 4A shows the phytosterol consumption and the AD production over time by this strain. Remarkably, the new strain was able to produce up to 7 g/L of AD. In these conditions, a

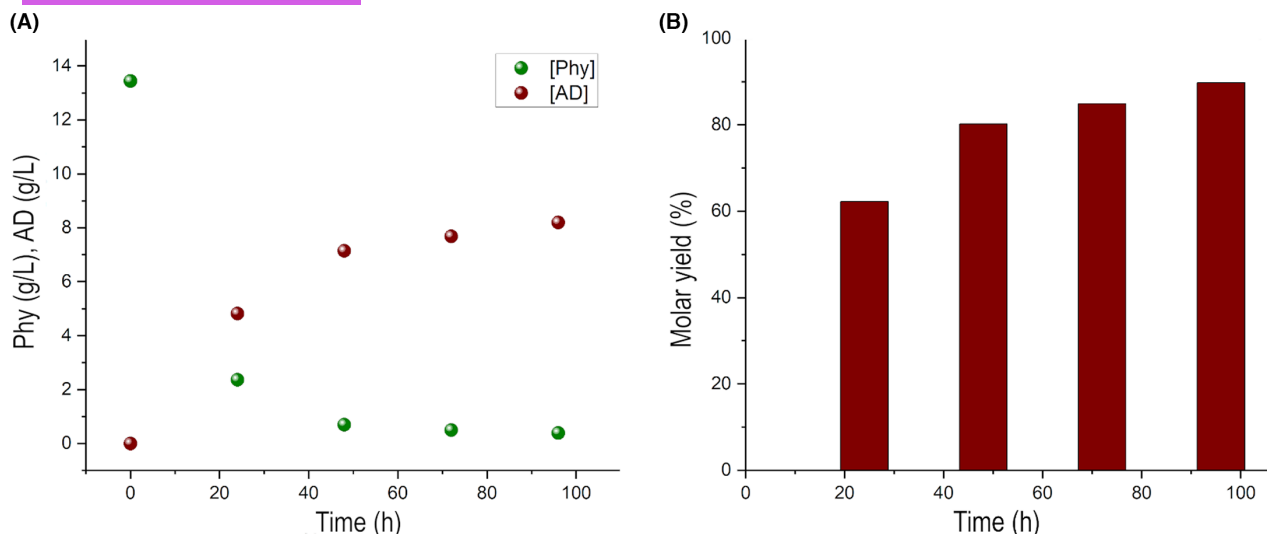


FIGURE 4 (A) 4-Androstene-3,17-dione (AD) production and phytosterols consumption over time of the MS6039-5941-6561 strain cultured in a bioreactor containing phytosterols at 10 g/L. (B) Molar yield of AD production of the MS6039-5941-6561 strain.

productivity of 0.17 g/L/h of AD is reached at 24 h and a productivity of 0.12 g/L/h at 48 h, indicating that the maximum production of AD occurs in the first 24 h. Moreover, molar yield of this biotransformation after 72 h reached 98% (Figure 4B). After 96 h we have not detected the presence of 4-HBC and only traces of non-consumed stigmasterol can be detected by GC–MS (Figure S3). The aliphatic side-chain of stigmasterol contains a double bond that might explain why this compound is metabolized less efficiently than cholesterol, campesterol or β -sitosterol. In this sense, β -oxidation of side-chain of cholesterol or phytosterols is not identical (Fujimoto et al., 1982; Wilbrink et al., 2012). Some biochemical steps are described to be specific for phytosterols side-chain degradation that in turn are not essential for cholesterol, like the CoA activation by FadD19 and the retro-aldol elimination of propionyl CoA by Ltp3 and/or Ltp4 aldol lyases occurring in the first β -oxidation cycle (Wilbrink et al., 2012; Wrońska et al., 2016). However, further CoA activation steps are followed by two rounds of β -oxidation and formation of a C19 steroid via aldolytic cleavage are identical to cholesterol side-chain degradation. Thus, it is becoming clear that phytosterols are substrates that require a more specific arsenal of enzymes than cholesterol for the degradation of the branched side-chain. Therefore, deepening the knowledge of the enzymes involved in this process will allow us to build more efficient production strains.

CONCLUSION

The results presented in this work demonstrate that the *MSMEG_6561* gene encoding a putative aldolase is involved in the production of 4-HBC in the metabolism

of sterols in *M. smegmatis*. This gene is homologous to the *sal* gene of some bacteria able to degrade bile salts. In *M. smegmatis* the gene appears to form part of an operon involved in a β -oxidation process that is not necessary for cholesterol degradation.

Since 4-HBC cannot be used by *M. smegmatis* as a carbon and energy source, the presence of this end-product can only be explained because the blockage of the pathway generates an accumulation of CoA-intermediaries that are not efficiently processed by the sterol side-chain β -oxidation enzymes. When *M. smegmatis* is cultured on high concentrations of sterols, the accumulation of CoA-derivatives can be toxic for the cells since it reduces the availability of CoA for other processes hindering the bacterial growth. In these cases, thiolases or aldolases having a broad range substrate specificity can help to release CoA from the overproduced CoA-derivatives (Haapalainen et al., 2006). The release of CoA is concomitant to the production of acetyl-CoA that can be readily metabolized.

It is worth to notice that, in contrast to the metabolism of bile-salts, where the resulting aldehyde is oxidized to a carboxyl group by the aldehyde dehydrogenase Sad (Holert et al., 2013), in *M. smegmatis* the aldehyde is reduced to an hydroxylic group.

Although we cannot completely discard that other genes could play a similar role to that of *MSMEG_6561*, at least this gene appears to be the main responsible for the production of 4-HBC in *M. smegmatis* under the tested conditions and in the absence of an active *msOpccR* reductase (Figure 1). Interestingly, *sal* gene does not appear to play the same critical role in *M. neoaurum* since the deletion of *msopccR* was enough to prevent the production of 4-HBC (Peng et al., 2021), therefore the role of the homologous *sal* genes in other *Mycobacterium* should be further investigated.

AUTHOR CONTRIBUTIONS

Gabriel Hernández-Fernández: Formal analysis (supporting); investigation (equal); methodology (equal); writing – original draft (supporting). **Miguel G. Acedos:** Formal analysis (supporting); investigation (equal); methodology (equal); writing – review and editing (supporting). **José L. García:** Conceptualization (lead); formal analysis (lead); funding acquisition (lead); supervision (equal); writing – review and editing (lead). **Beatriz Galán:** Conceptualization (lead); funding acquisition (lead); investigation (supporting); supervision (lead); writing – review and editing (lead).

ACKNOWLEDGMENTS

This work was supported by grants from Spanish Ministry of Science and Innovation PID2021-125370OB-I00, PID2019-110612RB-I00 from the Ministry of Economy and Competitiveness of Spain and grant from the Community of Madrid and the Structural Funds from European Union (Ref: S2018/BAA-4532 (ALGATEC-CM)). The grant for one of the authors (GHF) with reference FPU21/05101 is gratefully recognized. The authors would like to thank Ana Valencia for the technical assistance.

CONFLICT OF INTEREST STATEMENT

The authors have no conflict of interest.

ORCID

José L. García  <https://orcid.org/0000-0002-9238-2485>
Beatriz Galán  <https://orcid.org/0000-0002-2596-6034>

REFERENCES

- Capyk, J.K., Kalscheuer, R., Stewart, G.R., Liu, J., Kwon, H., Zhao, R. et al. (2009) Mycobacterial cytochrome P450 125 (Cyp125) catalyzes the terminal hydroxylation of C27 steroids. *The Journal of Biological Chemistry*, 284, 35534–35542.
- Donova, M. (2021) Microbial steroid production technologies: current trends and prospects. *Microorganisms*, 10, 53.
- Feng, J., Wu, Q., Zhu, D. & Ma, Y. (2022) Biotransformation enables innovations toward green synthesis of steroidal pharmaceuticals. *ChemSusChem*, 15, e202102399.
- Fernández-Cabezón, L., Galán, B. & García, J.L. (2018) New insights on steroid biotechnology. *Frontiers in Microbiology*, 9, 958.
- Galán, B., Uhía, I., García-Fernández, E., Martínez, I., Bahillo, E., Fuente, J.L. et al. (2017) *Mycobacterium smegmatis* is a suitable cell factory for the production of steroidal synthons. *Microbial Biotechnology*, 10, 138–150.
- Fujimoto, Y., Chem, C.S., Szeleczky, Z., Di Tullio, D. & Sih, C.J. (1982) Microbial degradation of the phytosterol sidechain. I. Enzymatic conversion of 3-oxo-24-ethylcholest-4-en-26-oic acid into 3-oxocholesterol-4-en-24-oic-acid and androst-4-ene-3,17-dione. *Journal of the American Chemical Society*, 104, 4718–4720.
- García, J.L., Uhía, I. & Galán, B. (2012) Catabolism and biotechnological applications of cholesterol degrading bacteria. *Microbial Biotechnology*, 5, 679–699.
- Gomila, M., Mulet, M., García-Valdés, E. & Lalucat, J. (2022) Genome-based taxonomy of the genus *Stutzerimonas* and proposal of *S. frequens* sp. nov. and *S. degradans* sp. nov. and emended descriptions of *S. perfectomarina* and *S. chloritidis-mutans*. *Microorganisms*, 10, 1363.
- Haapalainen, A.M., Meriläinen, G. & Wierenga, R.K. (2006) The thiolase superfamily: condensing enzymes with diverse reaction specificities. *Trends in Biochemical Sciences*, 31, 64–71.
- Holert, J., Jagmann, N. & Philipp, B. (2013) The essential function of genes for a hydratase and an aldehyde dehydrogenase for growth of *Pseudomonas* sp. strain Chol1 with the steroid compound cholate indicates an aldolytic reaction step for deacetylation of the side chain. *Journal of Bacteriology*, 195, 3371–3380.
- Jackson, M., Camacho, L.R., Gicquel, B. & Guilhot, C. (2001) Gene replacement and transposon delivery using the negative selection marker *sacB*. In: Parish, T. & Stocker, N.G. (Eds.) *Mycobacterium tuberculosis. Protocols*, Vol. 54. Totowa, NJ: Humana Press, pp. 59–75.
- Kutney, J.P., Herrington, E.J. & Spassov, G. (2003) *Process for fermentation of phytosterols to androstadienedione*. Patent WO 2003/064674.
- Nesbitt, N.M., Yang, X., Fontán, P., Kolesnikova, I., Smith, I., Sampson, N.S. et al. (2010) A thiolase of *Mycobacterium tuberculosis* is required for virulence and production of androstenedione and androstadienedione from cholesterol. *Infection and Immunity*, 78, 275–282.
- Nunes, V.O., Vanzellotti, N.C., Fraga, J.L., Pessoa, F.L.P., Ferreira, T.F. & Amaral, P.F.F. (2022) Biotransformation of phytosterols into androstenedione – a technological prospecting study. *Molecules*, 27, 3164.
- Ouellet, H., Guan, S., Johnston, J.B., Chow, E.D., Kells, P.M., Burlingame, A.L. et al. (2010) *Mycobacterium tuberculosis* CYP125A1, a steroid C27 monooxygenase that detoxifies intracellularly generated cholest-4-en-3-one. *Molecular Microbiology*, 77, 730–742.
- Pandey, A.K. & Sassetti, C.M. (2008) Mycobacterial persistence requires the utilization of host cholesterol. *Proceedings of the National Academy of Sciences of the United States of America*, 105, 4376–4380.
- Peng, H., Wang, Y., Jiang, K., Chen, X., Zhang, W., Zhang, Y. et al. (2021) Dual role reductase from phytosterols catabolism enables the efficient production of valuable steroid precursors. *Angewandte Chemie (International Ed. in English)*, 60, 5414–5420.
- Snapper, S.B., Melton, R.E., Mustafa, S., Kieser, T. & Jacobs, W.R., Jr. (1990) Isolation and characterization of efficient plasmid transformation mutants of *Mycobacterium smegmatis*. *Molecular Microbiology*, 4, 1911–1919.
- Song, S., He, J., Gao, M., Huang, Y., Cheng, X. & Su, Z. (2023) Loop pathways are responsible for tuning the accumulation of C19- and C22-sterol intermediates in the mycobacterial phytosterol degradation pathway. *Microbial Cell Factories*, 22, 19.
- Stover, C.K., de la Cruz, V.F., Fuerst, T.R., Burlein, J.E., Benson, L.A., Bennett, L.T. et al. (1991) New use of BCG for recombinant vaccines. *Nature*, 351, 456–460.
- Uhía, I., Galán, B., Kendall, S.L., Stoker, N.G. & García, J.L. (2012) Cholesterol metabolism in *Mycobacterium smegmatis*. *Environmental Microbiology Reports*, 4, 168–182.
- Uhía, I., Galán, B., Morales, V. & García, J.L. (2011) Initial step in the catabolism of cholesterol by *Mycobacterium smegmatis* mc² 155. *Environmental Microbiology*, 13, 943–959.
- Wang, X.X., Ke, X., Liu, Z.Q. & Zheng, Y.G. (2022) Rational development of mycobacteria cell factory for advancing the steroid biomanufacturing. *World Journal of Microbiology and Biotechnology*, 38, 191.
- Wilbrink, M.H., van der Geize, R. & Dijkhuizen, L. (2012) Molecular characterization of *ltp3* and *ltp4*, essential for C24-branched chain sterol-side-chain degradation in *Rhodococcus rhodochrous* DSM 43269. *Microbiology*, 158, 3054–3062.
- Wrońska, N., Brzostek, A., Szewczyk, R., Soboń, A., Dziadek, J. & Lisowska, K. (2016) The role of *fadD19* and *echA19* in sterol side chain degradation by *Mycobacterium smegmatis*. *Molecules*, 21, 598.



- Xu, L., Q., Liu, Y.J., Yao, K., Liu, H.H., Tao, X.Y. et al. (2016) Unraveling and engineering the production of 23, 24-bisnorcholesterol in sterol metabolism. *Scientific Reports*, 6, 1–13.
- Zhao, A., Zhang, X., Li, Y., Wang, Z., Lv, Y., Liu, J. et al. (2021) *Mycobacterium* cell factory for the production of steroid-based drug intermediates. *Biotechnology Advances*, 53, 107860.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Hernández-Fernández, G., Acedos, M.G., García, J.L. & Galán, B. (2023) Identification of the aldolase responsible for the production of 22-hydroxy-23,24-bisnorcholesterol-4-ene-3-one from natural sterols in *Mycobacterium smegmatis*. *Microbial Biotechnology*, 00, 1–10. Available from: <https://doi.org/10.1111/1751-7915.14270>

ANEXO II. Publicación 2

RESEARCH ARTICLE

Improving the production of 22-hydroxy-23,24-bisnorcholesterol-4-ene-3-one in *Mycolicibacterium smegmatis*

Gabriel Hernández-Fernández¹ | Miguel G. Acedos²  | Isabel de la Torre¹ | Juan Ibero¹ | José L. García¹  | Beatriz Galán¹ 

¹Department of Microbial and Plant Biotechnology, Centro de Investigaciones Biológicas Margarita Salas (CSIC), Madrid, Spain

²Advanced Biofuels and Bioproducts Unit, Department of Energy, Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas (CIEMAT), Madrid, Spain

Correspondence

Beatriz Galán, Department of Microbial and Plant Biotechnology, Centro de Investigaciones Biológicas Margarita Salas (CSIC), Madrid, Spain.
Email: bgalan@cib.csic.es

Funding information

Grant/Award Number: FPU21/05101 and PID2021-125370OB-I00; Ministerio de Economía y Competitividad, Grant/Award Number: PID2019-110612RB-I00; Comunidad de Madrid, Grant/Award Number: S2018/BAA-4532

Abstract

The 22-hydroxy-23,24-bisnorcholesterol-4-ene-3-one (4-HBC) is a C22 steroid synthesis of pharmaceutical interest that can be produced as a lateral end-product of the catabolism of natural sterols (e.g., cholesterol or phytosterols). This work studies the role of an aldehyde dehydrogenase coded by the *MSMEG_6563* gene of *Mycolicibacterium smegmatis*, named msRed, in 4-HBC production. This gene is located contiguously to the *MSMEG_6561* encoding the aldolase msSal which catalyses the retroaldol elimination of acetyl-CoA of the metabolite intermediate 22-hydroxy-3-oxo-cholesterol-4-ene-24-carboxyl-CoA to deliver 3-oxo-4-pregnene-20-carboxyl aldehyde (3-OPA). We have demonstrated that msRed reduces 3-OPA to 4-HBC. Moreover, the role of msOpccR reductase encoded by *MSMEG_1623* was also explored confirming that it also performs the reduction of 3-OPA into 4-HBC, but less efficiently than msRed. To obtain a *M. smegmatis* 4-HBC producer strain we deleted *MSMEG_5903* (*hsd4A*) gene in strain MS6039-5941 ($\Delta kshB1$, $\Delta kstD1$) that produces 4-androstene-3,17-dione (AD) from natural sterols (cholesterol or phytosterols). The triple MS6039-5941-5903 mutant was able to produce 9 g/L of 4-HBC from 14 g/L of phytosterols in 2 L bioreactor, showing a productivity of 0.140 g/L h⁻¹. To improve the metabolic flux of sterols towards the production of 4-HBC we have cloned and overexpressed the msSal and msRed enzymes in the MS6039-5941-5903 mutant rendering a production titer of 12.7 g/L with a productivity of 0.185 g/L h⁻¹, and demonstrating that the new recombinant strain has a great potential for its industrial application.

INTRODUCTION

Currently, the production of valuable steroid precursors for the synthesis of steroid pharmaceutical drugs is based on the utilization of mycobacterial strains that transform natural sterols into useful intermediates or synthons as building blocks, such as C19 and C22 steroids (Donova, 2021; Fernández-Cabezón et al., 2018; Nunes et al., 2022; Wang et al., 2024). At present, three main C19 steroid synthons are produced at industrial scale by

different mycobacterial strains using phytosterols as raw materials, i.e., 4-androstene-3,17-dione (AD), 1,4-androstadiene-3,17-dione (ADD) and 9 α -hydroxy-4-androstenedione (9-OH-AD). Whereas the C17 side-chain in C19 synthons is absent, C22 synthons are metabolites resulting from an incomplete degradation of the C17 side-chain, i.e., 22-hydroxy-23,24-bisnorcholesterol-1,4-dien-3-one (1,4-HBC), 22-hydroxy-23,24-bisnorcholesterol-4-ene-3-one (4-HBC) and 9,22-dihydroxy-23,24-bisnorcholesterol-4-ene-3-one (9-OH-4-HBC) (Figure 1) (Donova, 2018; Peng et al., 2021). C19

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs](https://creativecommons.org/licenses/by-nc-nd/4.0/) License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2024 The Author(s). *Microbial Biotechnology* published by John Wiley & Sons Ltd.

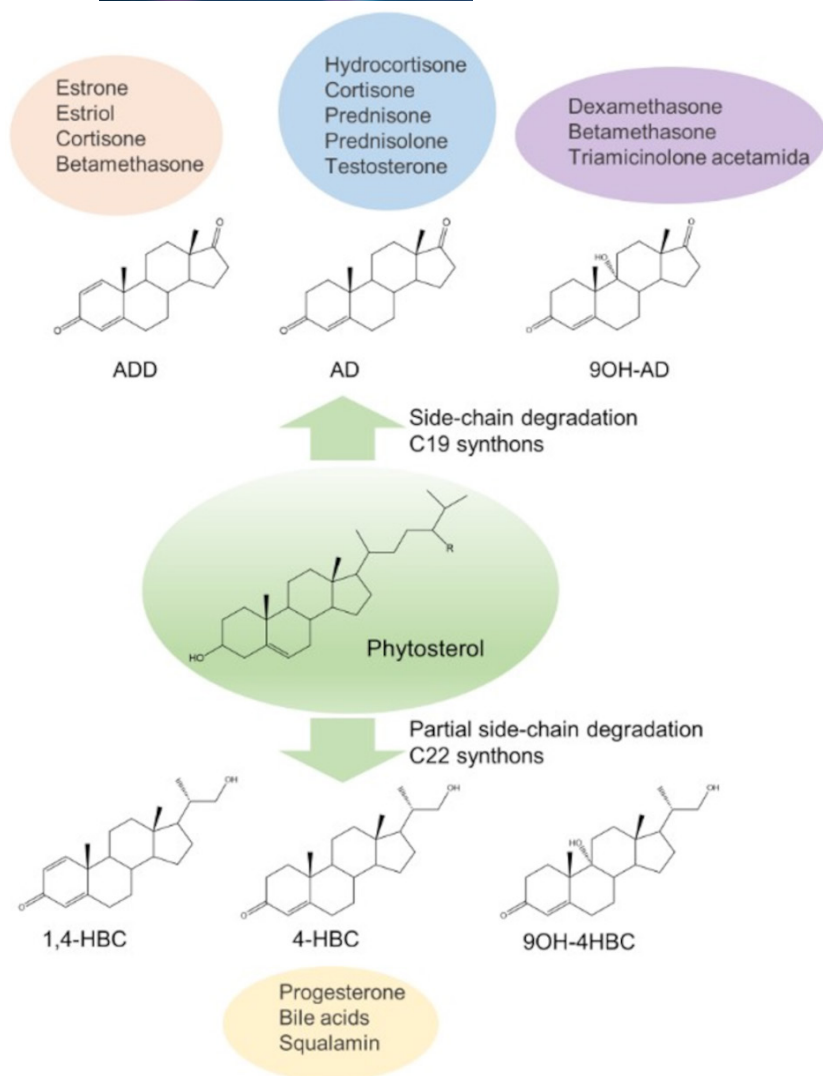


FIGURE 1 Chemical structure of relevant synthons biotechnologically produced from phytosterols.

synthons are used to synthesize testosterone, oestrone, oestradiol, hydrocortisone, methylprednisolone, betamethasone, or dexamethasone, among others, whereas C22 synthons are used to produce progesterone, obeticholic acid, ursodeoxycholic acid, chenodeoxycholic acid and lithocholic acid (Figure 1). Moreover, C22 synthons could be used to produce squalamin, which is currently in clinical trials for the treatment of macular degeneration (Fabre et al., 2022; Feng et al., 2022; Zhao et al., 2021).

The aerobic sterol degradation pathway of mycobacterial strains has been studied extensively and most of the key intermediates have been identified. Briefly, sterols are uptake by the transport system encoded by the *mce* operon (García-Fernández et al., 2017) and immediately oxidized to 5-cholesten-3-one and isomerized to 4-cholesten-3-one by a 3 β -hydroxysteroid dehydrogenase/isomerase (3 β -HSD) (Uhía et al., 2011). Simultaneously, the aliphatic chain is degraded following a mechanism similar to the β -oxidation of fatty acids, consisting of three sequential β -cyclic oxidations (Capyk et al., 2009; Ouellet

et al., 2010). Each cycle comprises enzymes such as ligases, dehydrogenases, hydratases and one aldolase complex responsible for the AD formation and the release of two propionyl-CoA molecules and one acetyl-CoA molecule (Figure 2) (García et al., 2012; Hernández-Fernández et al., 2023). Considering this information, it has been demonstrated that the knockout mutants of the coding genes of the 3-ketosteroid- Δ^1 -dehydrogenase KstD (MSMEG_5941), 3-ketosteroid-9 α -hydroxylase KshA (MSMEG_5925) and the KshA ferredoxin reductase component KshB (MSMEG_6039) of *Mycobacterium smegmatis* in different combinations allow the production of AD, ADD and 9-OH-AD (Fernández-Cabezón et al., 2018; Wang et al., 2022), while additional knockout of the *hsd4A* (MSMEG_5903) gene coding a 17 β -hydroxysteroid dehydrogenase lead to the accumulation of 4-HBC, 1,4-HBC and 9-OH-4-HBC (Liu et al., 2023; Peng et al., 2021; Xu et al., 2016; Zhao et al., 2023).

Although deeply studied, there are still unknown aspects of *Mycobacterium* steroid metabolism, and

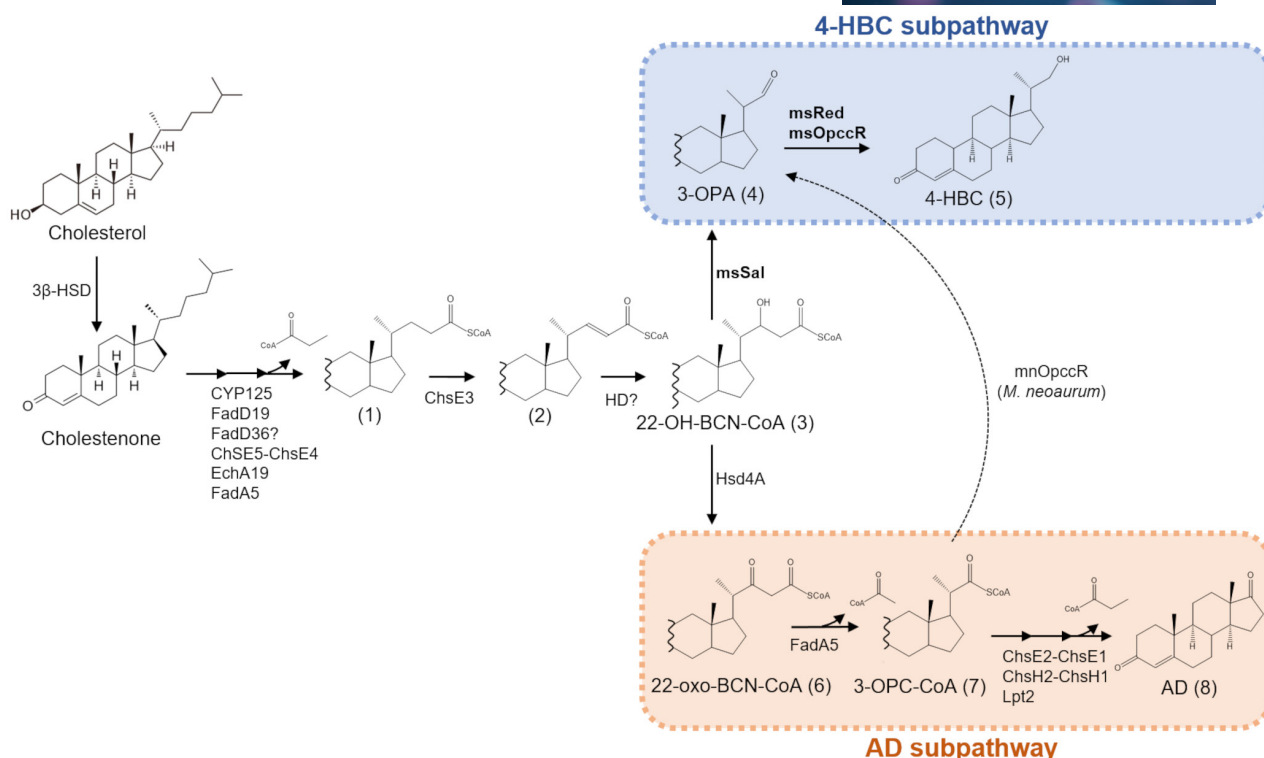


FIGURE 2 Cholesterol side-chain degradation pathway in *M. smegmatis*. The 4-HBC and the AD sub-pathways are indicated in blue and orange, respectively. Enzymes with a question mark are unknown. HD, hydratase. (1) 3-oxochol-4-en-24-oyl-CoA, (2) 3-oxochol-4,22-dien-24-oyl-CoA, (3) 22-hydroxyl-3-oxochol-4-en-24-oyl-CoA, (4) 3-oxo-4-pregnene-20-carboxyl aldehyde, (5) 22-hydroxy-23,24-bischolesterol-4-en-3-one, (6) 3,22-dioxochol-4-en-24-oyl-CoA, (7) 3-oxo-23,24-bischolesterol-4-en-22-oyl-CoA, (8) androstenedione.

in particular of *M. smegmatis*, that remains to be elucidated and are essential to achieve efficient industrial production of C19 and C22 synthons. For instance, the *kstD* and *kshB* knockout strain constructed to accumulate AD also accumulates a certain amount of 4-HBC as by-product, reducing the biotransformation yield and hindering the purification of AD (Galán et al., 2017). In this sense, Peng et al. (2021) have demonstrated a pivotal biosynthetic mechanism of the phytosterol catabolism that explained the presence of 4-HBC as a by-product in *Mycolicibacterium neoaurum*. They deciphered the dual role of the reductase mnOpccR in the phytosterols catabolism, which participates in two different metabolic branches to produce the key intermediate 4-HBC through a 4-e reduction of 3-oxo-4-pregnene-20-carboxyl-CoA (3-OPC-CoA) to 3-oxo-4-pregnene-20-carboxyl aldehyde (3-OPA) in the C-terminal domain followed by a 2-e reduction of 3-OPA to 4-HBC performed by the N-terminal domain of mnOpccR (Figure 2). Inactivation of mnOpccR eliminated 4-HBC production and led to the exclusive production of AD, while overexpression of mnOpccR together with Hsd4A inactivation led to the production of 4-HBC from phytosterols. This study provided a very efficient production route of AD and 4-HBC synthons in *M. neoaurum*. However, unlike what happens in *M. neoaurum*, the C-terminal of the homologous protein in *M. smegmatis*

(msOpccR) is not active and does not catalyse the reduction of 3-OPC-CoA to 3-OPA (Peng et al., 2021). This result suggested that other enzymes might exist in *M. smegmatis* to perform this task.

We have recently demonstrated that the enzyme responsible for the conversion of 22-hydroxy-3-oxo-cholest-4-ene-24-carboxyl-CoA (22-OH-BCN-CoA) into 3-OPA in *M. smegmatis* is a Sal-like aldolase encoded by *MSMEG_6561* gene that catalyses the retroaldol elimination of acetyl-CoA (Hernández-Fernández et al., 2023). The retroaldol reaction was previously described in the degradation of the cholic acid side chain in *Sacharophagus degradans* and it is performed by the steroid aldolase Sal encoded by the *sal* gene (Holert et al., 2013). Sal-like aldolase from *M. smegmatis*, hereafter msSal, shows a 57% similarity with that of *S. degradans* (O0E10244.1) and both compounds 22-OH-BCN-CoA, precursor of 4-HBC, and cholic acid are structurally similar. The deletion of *MSMEG_6561* gene (*sal*) in *M. smegmatis* led to an increase in the production and an improvement in the yield of AD from natural sterols and avoided completely the production of 4-HBC as a by-product (Hernández-Fernández et al., 2023).

These data suggest that a further reduction of 3-OPA into 4-HBC in *M. smegmatis* should be achieved by the N-terminal domain of msOpccR reductase or by another unidentified reductase. In this work we demonstrate

that 3-OPA is reduced into 4-HBC by an aldehyde dehydrogenase encoded by the *MSMEG_6563* gene located contiguously to the Sal-like aldolase. Moreover, we have also demonstrated that the msOpccR enzyme encoded by *MSMEG_1623* gene is active in *M. smegmatis*, but performs a less efficient reduction.

EXPERIMENTAL PROCEDURES

Chemicals

Steroids used in this work are 4-androstene-3,17-dione (AD), purchased from TCI America, cholesterol and cholestenone from Sigma (Germany), 4-HBC from Clearsynth (India) and phytosterols from Gadea Biopharma (Spain), which are conformed by β -sitosterol (54.4%), campesterol (36.8%) stigmasterol (8.7%) and brassicasterol (2.16%) (w/w). Chloroform, acetonitrile, glycerol and methanol were purchased from Merck (Darmstadt, Germany). N,O-bis(trimethylsilyl) trifluoroacetamide (BSTFA), pyridine, Tween80 and tyloxapol were purchased from Sigma (Germany). Oligonucleotides were from IDT Technologies.

Bacterial strains and culture conditions

All strains, plasmids and primers used are grouped in Table 1. *M. smegmatis* mc² 155 and the subsequent mutants were grown in LB medium with 0.05% Tween80 as rich medium at 37°C and with orbital shaking at 200 rpm. *Escherichia coli* strains were also grown in LB medium at 37°C with orbital shaking at 200 rpm. As solid media, LB agar plates and 7H10 agar (Difco) with sucrose 10% (w/v) were used for *E. coli* and *M. smegmatis*. Regarding the biotransformation assays, minimal medium Middlebrook 7H9 broth (Difco) was used supplemented with 9 mM glycerol and 0.4 g/L cholesterol. The cholesterol stock was previously prepared at 5 mM in tyloxapol 10% (v/v) and it was sonicated for 30 min and autoclaved twice until its total dissolution. A *M. smegmatis* preinoculum was first grown during 48 h and washed with saline solution 0.85% containing 0.05% Tween80 was used to inoculate 25 mL of biotransformation media at OD₆₀₀ = 0.05. Cultures were then grown at 200 rpm and 37°C for 96 h and 500 μ L samples were taken every 24 h and kept at -20°C until their analysis.

Bioreactor experiments were performed as described before (Hernández-Fernández et al., 2023). Briefly, culture medium contained corn steep liquor (20 g/L), sodium nitrate (5.4 g/L), ammonium chloride (0.28 g/L) and monobasic sodium phosphate (0.28 g/L). 14 g of phytosterols were dissolved by autoclaving in 160 mL of sunflower oil, and mixed with 740 mL of the culture medium, while the inoculum volume was 10%.

The biotransformation was carried out in a working volume of 1 L in a commercial stirred tank bioreactor (STBR) of 2 L total volume (Sartorius AG, Germany). Process temperature was maintained at 37°C, pH was adjusted to 7.4, but it was not controlled during the process. Stirrer speed was set at 600 rpm and an air flow rate of 1 vvm was used. Dissolved oxygen was maintained above 20% throughout the process. A 15 mL sample was taken every 24 h and kept at -20°C. Different antibiotics were used when needed, being their final concentration 5 μ g/mL of gentamicin and 20 μ g/mL of kanamycin for *M. smegmatis* strains, while 10 μ g/mL of gentamicin and 50 μ g/ μ L of kanamycin were used for the *E. coli* strains.

Construction of knock-out mutant strains

The *MSMEG_5903* gene deletion in the *M. smegmatis* MS6039-5941 strain was achieved by homologous recombination with pJQ200x plasmid as described previously (Hernández-Fernández et al., 2023). Two fragments of about 600 bp containing the upstream and downstream regions of *MSMEG_5903* were amplified by PCR using the primers listed in Table 1. The genomic DNA from *M. smegmatis* was obtained as described (Uhía et al., 2011). The two fragments amplified were digested with the corresponding enzyme and ligated with the pJQ200x backbone. *E. coli* DH10B chemical competent cells (Invitrogen) were used. This cloning led to the pJQ5903 plasmid that was then electroporated into competent MS6039-5941. *M. smegmatis* electrocompetent cells were obtained as follows: a 100 mL culture of LB with 0.05% Tween80 was inoculated at an OD₆₀₀ 0.01 and grown over-night (~18 h). When cells reached an OD₆₀₀ 0.8–1.0, they were harvested and washed three times with glycerol 10% (v/v) and 0.05% Tween80. Finally, the pellet was resuspended in 1 mL of the same solution, aliquoted in 200 μ L and kept at -80°C until used. Electroporation conditions used were 2.5 kv, 25 μ F y 1000 Ω in a Gene pulser/Pulse controller electroporator (Bio-Rad). Cells with a single recombination event were selected on 7H10 agar plates containing gentamycin. Insertion of the pJQ5903 plasmid was assessed by spreading catechol over the colonies, since the *xyIE* gene present in the plasmid turns yellow in the bacterial colonies. Colonies were then contra-selected in 10% (w/v) sucrose in order to select for double cross-overs. A single colony was grown in 10 mL of LB medium with 5 μ g/mL gentamicin up to OD₆₀₀ of about 0.8 and 200 μ L of a 1:100 dilution was plated onto 7H10 agar plates with 10% sucrose. The sucrose-resistant colonies that have undergone a potential double cross-over were screened for gentamicin sensitivity. The mutant strain MS6039-5941-5903 was finally analysed by PCR and DNA sequencing to confirm the deletion of *MSMEG_5903* gene.

TABLE 1 Stains, plasmids and oligonucleotides used in this work.

Strains	Genotype and/or description	Reference
<i>M. smegmatis</i>		
mc ² 155	<i>ept-1</i> , mc ² 6 mutant efficient for electroporation	Snapper et al., 1990
MS6039-5941	<i>M. smegmatis</i> mc ² 155 Δ MSMEG_6039 (Δ kshB1) and Δ MSMEG_5941 (Δ kstD1)	Galán et al., 2017
MS6039-5941-5903	<i>M. smegmatis</i> mc ² 155 Δ MSMEG_6039 (Δ kshB1), Δ MSMEG_5941 (Δ kstD1) and Δ MSMEG_5603 (Δ hsd4A)	This work
MS6039-5941-5903 (pMV261)	<i>M. smegmatis</i> mc ² 155 Δ MSMEG_6039 (Δ kshB1), Δ MSMEG_5941 (Δ kstD1) and Δ MSMEG_5603 (Δ hsd4A) with pMV261	This work
MS6039-5941-5903 (pMVmsSal)	<i>M. smegmatis</i> mc2155 Δ MSMEG_6039 (Δ kshB1), Δ MSMEG_5941 (Δ kstD1) and Δ MSMEG_5603 (Δ hsd4A) with pMVmsSal	This work
MS6039-5941-5903 (pMVmsSal-msRed)	<i>M. smegmatis</i> mc ² 155 Δ MSMEG_6039 (Δ kshB1), Δ MSMEG_5941 (Δ kstD1) and Δ MSMEG_5603 (Δ hsd4A) with pMVmsSal-msRed	This work
MS6039-5941-5903 (pMVmsSal-msOpccR)	<i>M. smegmatis</i> mc ² 155 Δ MSMEG_6039 (Δ kshB1), Δ MSMEG_5941 (Δ kstD1) and Δ MSMEG_5603 (Δ hsd4A) with pMVmsSal-msOpccR	This work
<i>E. coli</i>		
DH10B	F ⁻ , <i>mcrA</i> , 1(<i>mrr hsdRMS-mcrBC</i>), 880 Δ lacZ1M15, Δ lacX74, <i>deoR</i> , <i>recA1</i> , <i>araD139</i> , Δ (<i>ara-leu</i>)7697, <i>galU</i> , <i>galK</i> , λ^- , <i>rpsL</i> , <i>endA1</i> , <i>nupG</i>	Invitrogen
DH10B (pMVmsSal)	<i>E. coli</i> DH10B with pMVmsSal	Hernández-Fernández et al., 2023
DH10B (pMVmsSal-msRed)	<i>E. coli</i> DH10B with pMVmsSal-msRed	This work
DH10B (pMVmsSal-msOpccR)	<i>E. coli</i> DH10B with pMVmsSal-msOpccR	This work
DH10B (pJQ5903)	<i>E. coli</i> DH10B with pJQ5903	This work
Plasmids	Description	Reference
pJQ200x	Suicide vector used to perform directed mutagenesis in <i>Mycobacterium</i> , P15A <i>ori</i> , <i>sacB</i> , Gm ^r , <i>xyIE</i>	Jackson et al., 2001
pJQ5903	pJQ200x containing one fragment upstream and another downstream of MSMEG_5903 between <i>Xba</i> I- <i>Sac</i> I	This work
pMV261	<i>Mycobacterium/E. coli</i> shuttle vector with the kanamycin resistance <i>aph</i> and the promoter from the <i>hsp60</i> gene from <i>M. tuberculosis</i>	Stover et al., 1991
pMVmsSal	pMV261 harbouring the MSMEG_6561 gene encoding Sal from <i>M. smegmatis</i> mc2 155 between <i>Bam</i> HI- <i>Hind</i> III	Hernández-Fernández et al., 2023
pMVmsSal-msRed	pMV261 harbouring the MSMEG_6561 gene encoding Sal between <i>Bam</i> HI- <i>Hind</i> III and the MSMEG_6563 gene between <i>Hind</i> III- <i>Hpa</i> I from <i>M. smegmatis</i> mc2 155	This work
pMVmsSal-msOpccR	pMV261 harbouring the MSMEG_6561 gene encoding Sal between <i>Bam</i> HI- <i>Hind</i> III and the MSMEG_1623 gene between <i>Hind</i> III- <i>Hpa</i> I from <i>M. smegmatis</i> mc2 155	This work
Primers	Description	Sequence
MSMEG_5903_UP_Fw	Amplification of the upstream fragment of MSMEG_5903 with <i>Xho</i> I and <i>Avr</i> II	CTAGCTCGAGCTCACGGATCGTGCGGCTCC
MSMEG_5903_UP_Rv		CTAGCCTAGGCGGCGACTGCACTGATGGAG
MSMEG_5903_DOWN_Fw	Amplification of the downstream fragment of MSMEG_5903 with <i>Avr</i> II and <i>Xho</i> I	CTAGCCTAGGCGGACAGGTGAGATCGGC
MSMEG_5903_DOWN_Rv		CTAGCTCGAGCATCTGCGCGACGGTCTCGC
MSMEG_6563_Fw	Amplification of MSMEG_6563 including an RBS and <i>Hind</i> III- <i>Hpa</i> I restriction sites	ACCAAGCTTGACCTAAGGAGGTGAATATGACA GCGACGA
MSMEG_6563_Rv	Amplification of MSMEG_1623 including an RBS and <i>Hind</i> III- <i>Hpa</i> I restriction sites	TAGGATGTTAACTCATGTACGCCTCCTCGACGC
MSMEG_1623_Fw		CTCAAGCTTGACCTAAGGAGGTGAATATGCGC TATGTCGTT
MSMEG_1623_Rv		CGCTGTTAACACGACTACCAGTGACGCGCGG

Construction of pMVmsSal-msRed and pMVmsSal-msOpccR plasmids

MSMEG_6563 and MSMEG_1623 genes were amplified by PCR using primers indicated in Table 1. They were then digested with *Hind*III and *Hpa*I restriction enzymes and cloned into pMVmsSal (Hernández-Fernández et al., 2023). Thus, the plasmids pMVmsSal-msRed and pMVmsSal-msOpccR were generated and transformed in *E. coli* DH10B. Plasmids were sequenced and transformed in electrocompetent cells of *M. smegmatis* resulting in the MS6039-5941-5903 (pMVmsSal-msRed) and MS6039-5941-5903 (pMVmsSal-msOpccR) recombinant strains.

Chromatographic analysis

A GC/MS analysis was performed in order to identify and quantify the different steroids as described (Hernández-Fernández et al., 2023). Samples of 200 µL of each culture were extracted twice with twice the volume of chloroform. Previously to its extraction, 25 µL of a solution of 10 mM cholestenone dissolved in chloroform were added to the aliquots as internal standards (ISTD). The chloroform fraction was concentrated by evaporation overnight at RT and trimethylsilyl ether derivatives were formed by reaction with 50 µL of TMT-BSTFA and 50 µL of pyridine and heating at 80 °C for 20 min. Steroid standards were derivatized in the same way and used to calculate calibration curves for steroid quantification (Figure S1). An Agilent 7890A gas chromatograph equipped with a 30 m × 0.25 mm i.d. capillary column (0.25 µm film thickness) HP-5MS (5% diphenyl 95% dimethylpolysiloxane from Agilent Technologies) and coupled to an Agilent 5975C mass detector (Agilent Technologies, USA) was used for the analysis. Mass spectra were recorded in electron impact (EI) mode at 70 eV within the *m/z* range 50–550 and working conditions for each sample were as follows: split ratio (20:1), injector temperature, 320 °C; column temperature 240 °C for 3 min, then heated to 320 °C at 5 °C/min. For quantification of the peak area, the quantitative masses were 329 + 458 *m/z* for cholesterol; 343 + 384 *m/z* for cholestenone; 244 + 286 *m/z* for AD and 270 + 330 *m/z* for 4-HBC in selected ions of monitoring. EI mass spectra and retention data were compared to those present in the NIST Mass Spectral Database and commercial standards (NIST 2011) in order to assess the identity of the different steroids.

Bioreactor samples were analysed by HPLC-MS. First, samples were centrifuged at RT and 3800 rpm in a 5810 R centrifuge (Eppendorf) for 30 min. 50 µL of the oil phase were dissolved in 1 mL of acetonitrile as well as the calibration curve of AD and 4-HBC, ranging from 0.5 to 15 g/L (Figure S2). Then, 20 µL of the samples were analysed using a high-efficiency liquid chromatography coupled to a photodiode array detector and mass spectroscopy. The equipment used was 1260 Infinity II

(Agilent) with an InfinityLab LC/MSD XT (Agilent), presenting an automatic injector, a diode array detector and an ion trap with electrospray. The column used was Poroshell 120 EC-C18 (4 µm; 4.6 × 100 mm) (Agilent). The separation method consisted of a flow rate of 1 mL/min with 70% of solvent A (H₂O + 0.1% formic acid) and 30% of solvent B (acetonitrile + 0.1% formic acid) for 5 min. Solvent B was then increased to 55% for 20 min and it was again reduced to 30% for 2 min. Finally, the proportion of 70% of solvent A and 30% of solvent B was maintained for 3 min. The elution was monitored at 195, 210, 230, 250 and 275 nm. Steroid compounds were identified using the commercial compound comparing mass spectra and retention time.

Quantification of the initial phytosterol concentration in sunflower oil was assessed by HPLC as described (Yang et al., 2023). Serial dilutions of phytosterols provided by Gadea Biopharma (Spain) ranging from 0.5 to 15 g/L were prepared in 200 µL of ethyl acetate to calculate a calibration curve (Figure S2). Samples and sterol standards were analysed in a 1260 Infinity II LC System (Agilent) HPLC coupled to a photodiode array detector. The column used was Eclipse Plus C18 (5 µm; 4.6 × 250 mm) (Agilent) heated at 30 °C, the mobile phase was methanol: acetonitrile (80:20), the flow rate was set at 1 mL/min and 20 µL of each sample were injected. Each run lasted 30 min and the detection was performed by UV at 195, 205 and 210 nm.

Calculations

The product molar yield (MY) was determined according to the following equation (Equation 1):

$$MY (\%) = \frac{[\text{Synthon}] (\text{mM})}{[\text{Sc}] (\text{mM})} \times 100 \quad (1)$$

where the term “Synthon” refers to any C19 or C22 steroid and “Sc” refers to consumed sterols.

Productivity (*P*) of 4-HBC was determined according to the following equation (Equation 2):

$$P \left(\frac{\text{g}}{\text{L} \cdot \text{h}} \right) = \frac{[4 - \text{HBC}] (\text{g/L})}{t (\text{h})} \quad (2)$$

where the term “*t*” refers to time (h).

RESULTS AND DISCUSSION

Engineering *M. smegmatis* for 4-HBC production

Along with C19 steroids, C22 steroids can be used as precursors to produce progestational and adrenocortical hormones in the pharmaceutical industry. Xu et al. (2016) developed a mutant *M. neoaurum* strain,

named $\Delta\text{hsd}4\text{A}$, that could transform sterols into 4-HBC. *Hsd4A* encodes a 17-beta-hydroxysteroid dehydrogenase involved in the oxidation of 22-OH-BCN-CoA to 3,22-dioxo-cholest-4-ene-24-carboxyl-CoA (22-OBNC-CoA) that finally leads to the formation of AD (Figure 2). When *hsd4A* gene is deleted together with *kstD* and *kshAB* genes, the sterol degradation pathway was deviated to the accumulation of 4-HBC, although AD was not abolished completely (Peng et al., 2021).

M. smegmatis contains a Hsd4A enzyme encoded by *MSMEG_5903* gene located in one of the cholesterol degradation clusters of this strain, ranging from *MSMEG_5903* to *MSMEG_5943*. Therefore, to obtain a *M. smegmatis* 4-HBC-producing strain, we deleted *MSMEG_5903* in the AD-producer strain MS6039-5941 (ΔkshB1 , ΔkstD1) (Galán et al., 2017) by double homologous recombination, obtaining the triple mutant MS6039-5941-5903 strain. When growing in glycerol as the only carbon and energy source this strain showed a normal growth phenotype when compared to the parental strain (Figure 3A).

The production of 4-HBC by the mutant MS6039-5941-5903 was assessed in a biotransformation assay developed in flasks using 0.4 g/L of cholesterol as substrate and glycerol as carbon and energy source. Samples taken every 24 h and analysed by GC-MS confirmed that the triple mutant accumulated 4-HBC as the main product, reaching a final concentration

of 0.277 g/L (Figure 3A,C). However, AD was also detected in the biotransformation supernatants at a final concentration of 0.015 g/L (Figure 3B,C).

Considering that the contamination of AD observed in our mutant, as well as in other *Hsd4A* mutants (Peng et al., 2021; Xu et al., 2016), is not desirable during industrial production because it reduces the final yield and hinders the purification of the final product, increasing the production cost, we decided to construct a more efficient 4-HBC producer strains.

Identification of genes involved in 4-HBC production in *M. smegmatis*

In a previous work, we demonstrated that *msSal* coded by the *MSMEG_6561* gene initiates the 4-HBC production in *Mycobacterium* mutants by catalysing the retroaldol elimination of acetyl-CoA of the intermediate 22-OH-BCN-CoA to deliver 3-OPA (Figure 2). *MSMEG_6561* gene is part of a putative operon *MSMEG_6560-MSMEG_6563*, hereafter operon 4-HBC (Figure 4) spanning 4920 pb within the *M. smegmatis* genome including the following genes: *MSMEG_6560* (*msShy*), that codes a putative acyl-dehydratase homologous to *Shy* from *S. degradans* Chol1 (44.5% ID); *MSMEG_6562* codes a putative ACAD enzyme similar to *ChsE3* from *M. tuberculosis* strain H37Rv (36.4% ID) and *CasC* from *Rhodococcus jostii* sp. strain RHA1 (34.8% ID) that contain two fused

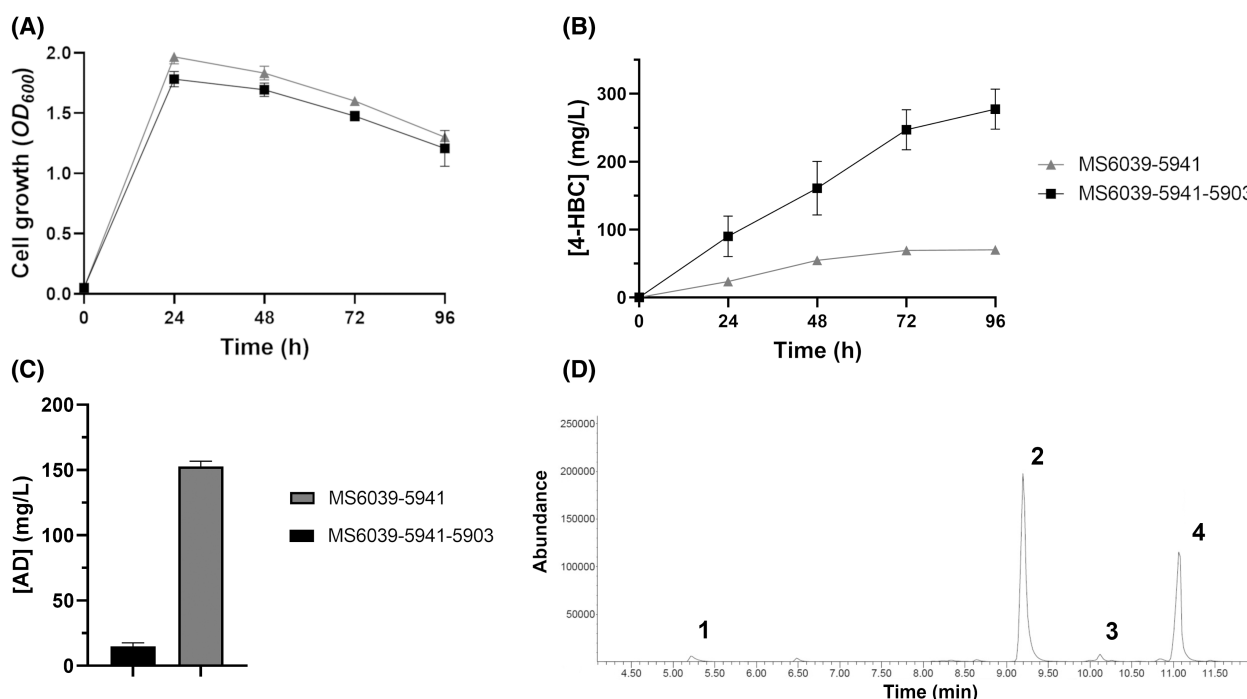


FIGURE 3 (A) Growth curve of strains MS6039-5941 and MS6039-5941-5903 in glycerol (B) Growth curve and 4-HBC production by strains MS6039-5941 and MS6039-5941-5903 from 0.4 g/L of cholesterol in flask. (C) AD accumulation by strains MS6039-5941 and MS6039-5941-5903 after 96 h of cholesterol biotransformation in flask. (D) Analysis by GC/MS of the products accumulated in the culture of MS6039-5941-5903 after 96 h with cholesterol at 0.4 g/L. 1, AD; 2, 4-HBC; 3, cholesterol; 4, cholestenone (ISTD).

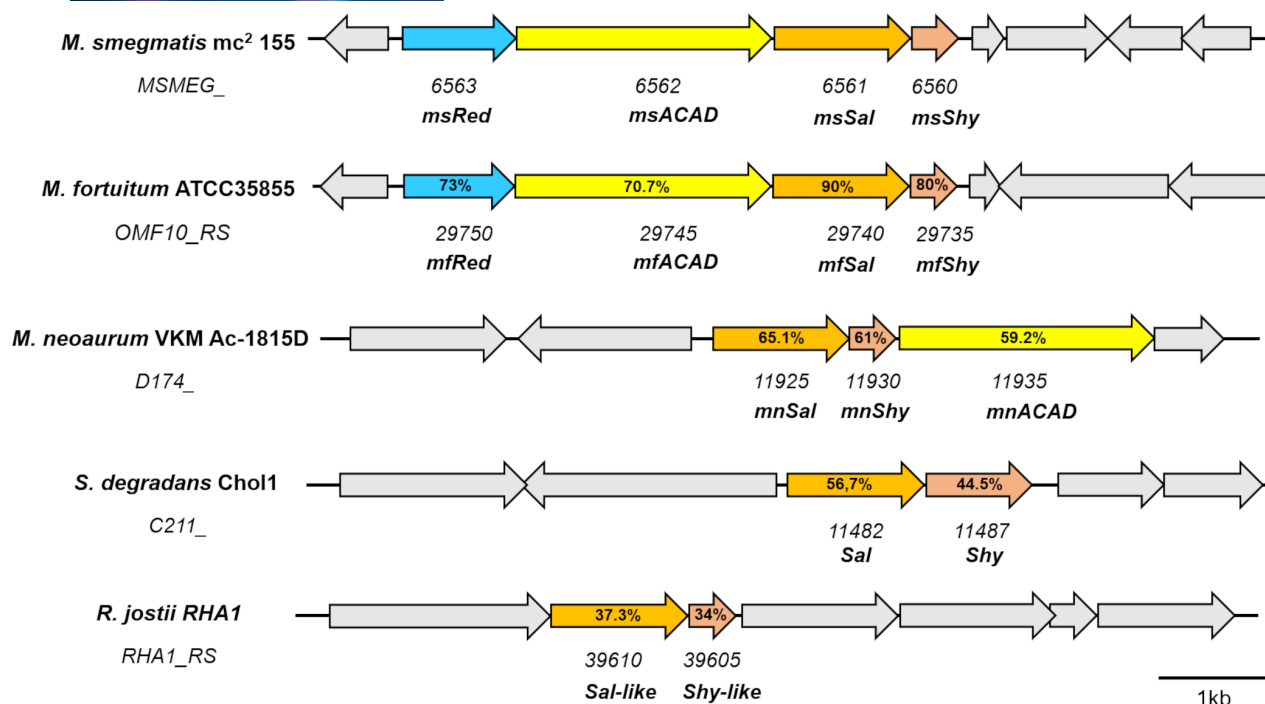


FIGURE 4 Genetic organization of the putative 4-HBC operon in *M. smegmatis* mc2 155 including msShy (MSMEG_6560), msSal (MSMEG_6561), msACAD (MSMEG_6562) and msRed (MSMEG_6563) genes. This cluster is compared with other genetic organizations of other cholate-degrading bacteria, showing the percentage of identity with the proteins encoded by the 4-HBC cluster.

acyl-CoA dehydrogenase domains in a single polypeptide and catalyse the dehydrogenation of 5-carbon CoA ester substituents on the D ring (Wipperman et al., 2014). This enzyme might participate in the C5-carboxylic side chain degradation catalysing a dehydrogenation reaction of compound 5 (Figure 5). Lastly, MSMEG_6563 is annotated as a dihydrokaempferol 4-reductase that contains a Rossmann-fold NAD(P)-binding domain that may function as an oxidoreductase that might be involved in the reaction of 3-OPA to 4-HBC. However, this enzyme hereafter msRed, is not homologous to the described NAD⁺-dependent steroid aldehyde dehydrogenase Sad performing this reaction in *Pseudomonas stutzeri* Chol1. Sad homologous are encoded within the bile acid degradation gene clusters of *Pseudomonas putida* DOC21, *Comamonas testosteroni*, and *Azoarcus* sp. strain Aa7 (Holert et al., 2013; Yücel et al., 2019), while no homologues of Sad exist in proximity to the cholic acid degradation gene cluster of *R. jostii* RHA1 or in *Sphingobium* sp. strain Chol1 genome and related bile-acid degrading Sphingomonadaceae (Figure 4) (Feller et al., 2021). A comparison of the genes responsible for C5-side chain degradation revealed significant similarities and differences between the genomes of *M. smegmatis* and *M. neoaurum* CCTCC AB2019054. These genomic comparisons were carried out using the available genome sequence of *M. neoaurum* VKM Ac-1815D since the genome of *M. neoaurum* CCTCC AB2019054 is not available. Nevertheless, the 17 gene sequences belonging to this strain already available in the databases are 100%

identical to those from VKM Ac-1815D suggesting that both strains are most probably identical.

We identified a genomic region in *M. neoaurum* VKM Ac-1815D spanning from D174_11925 to D174_11940 containing genes homologous to those of the genomic region MSMEG_6563-MSMEG_6560 from *M. smegmatis*. The VKM Ac-1815D region encodes a putative aldolase similar to msSal (65% ID), an acyl-CoA dehydrogenase similar to MSMEG_6562 (59% ID) and an acyl dehydratase similar to MSMEG_6560 (61% ID) (Figure 4). However, this genomic region does not contain a msRed reductase, since this reaction is catalysed by the mnOpccR reductase coded by the D174_07590 gene located far away from this genomic region. The recently identified HBC operon of *M. neoaurum* ATCC 25795 does not contain a homologous protein of msRed and it has been suggested that, in this strain, more than one specific enzyme is responsible for this activity (Zhao et al., 2023). We found that the HBC operon in *Mycobacterium fortuitum* ATCC35855 has the same gene structure as the one in *M. smegmatis* (Figure 4). Moreover, a gene encoding a protein homologous to msRed is present in *M. fortuitum* although its role is not known. The 4-HBC metabolism resembles the side chain degradation of bile acids described in β -proteobacteria that involves a previous hydration of (22E)7 α ,12 α -dihydroxy-3-oxochola-1,4,22E-triene-24-oyl-CoA (DHOCTO-CoA) (Compound 1) by the acyl-dehydratase Shy, forming 7 α ,12 α ,22-trihydroxy-3-oxochola-1,4-diene-24-oyl-CoA (THOCTO-CoA)

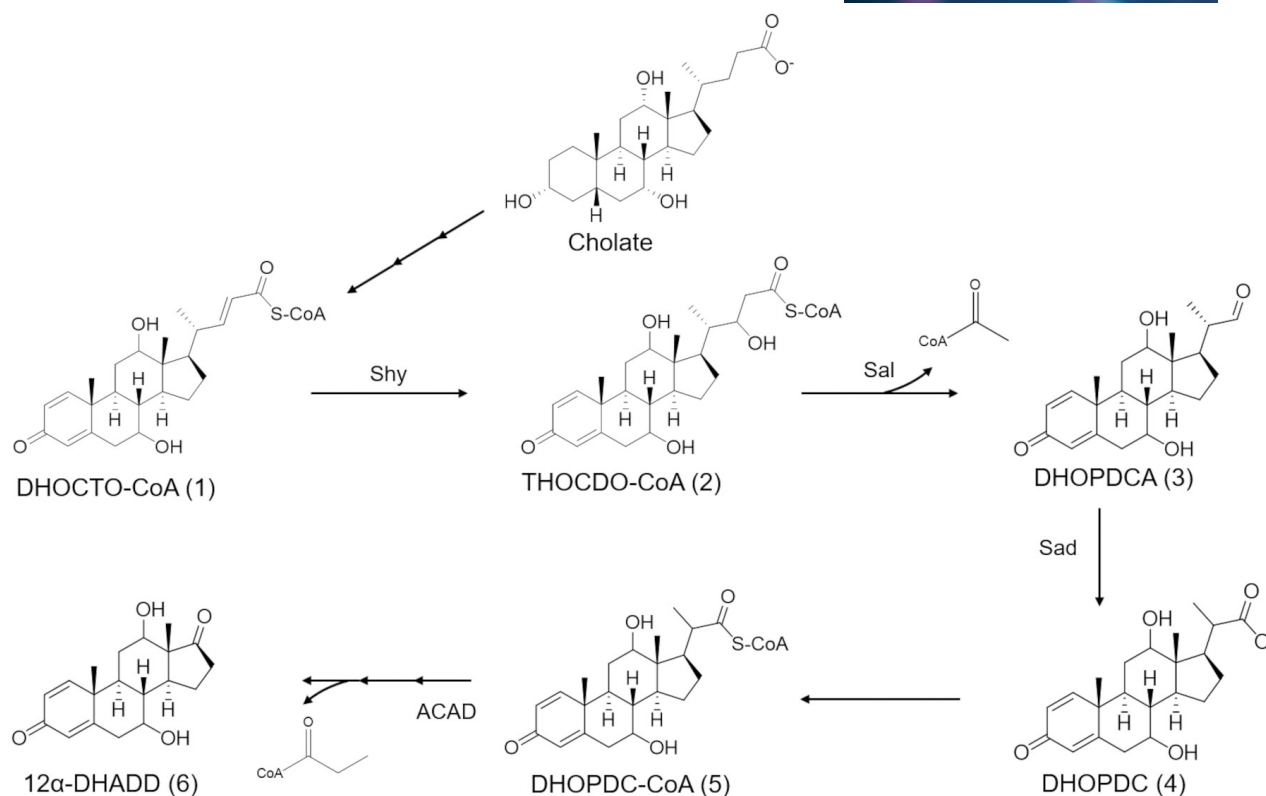


FIGURE 5 Proposed cholate side-chain degradation pathway in the cholate-degrading strain *S. degradans* Chol1. Shy: Steroid hydratase; Sal: Steroid aldolase; Sad: Steroid aldehyde dehydrogenase; ACAD: Acyl-CoA dehydrogenase. Compound 1, (22E)7 α ,12 α -dihydroxy-3-oxochola-1,4,22E-triene-24-oyl-CoA (DHOCTO-CoA); compound 2, 7 α ,12 α ,22-trihydroxy-3-oxochola-1,4-diene-24-oyl-CoA (THOCDO-CoA); compound 3, 7 α ,12 α -dihydroxy-3-oxopregna-1,4-diene-20S-carbaldehyde (DHOPDCA); compound 4, acid 7 α ,12 α -dihydroxy-3-oxopregna-1,4-diene-20-carboxylic (DHOPDC); compound 5, acid 7 α ,12 α -dihydroxy-3-oxopregna-1,4-diene-20-carboxyl-CoA (DHOPDC-CoA); compound 6, 7 α ,12 α -dihydroxy-androsta-1,4-diene-9,17-dione (12 α -DHADD).

(Compound 2) (Figure 5) (Holert et al., 2013). Then, a subsequent retroaldol elimination of acyl-CoA by the aldolase Sal occurs, delivering an aldehyde derivative 7 α ,12 α -dihydroxy-3-oxopregna-1,4-diene-20S-carbaldehyde (DHOPDCA) (Compound 3), that is transformed by the aldehyde dehydrogenase Sad into acid 7 α ,12 α -dihydroxy-3-oxopregna-1,4-diene-20-carboxylic (DHOPDC) (Compound 4). Then, DHOPDC is further activated in the presence of ATP and CoA by an acyl-CoA synthetase to DHOPDC-CoA to start a second round in which the C3-side-chain is oxidized at the C17-C20 bond by an acyl-CoA dehydrogenase (ACAD) to prepare the cleavage of the remaining C3 unit. Subsequently, the double bond is hydrated followed by the release of a propionyl-CoA moiety by the action of an unknown thiolase enzyme. MsShy, as its homologous Shy in the bile acid metabolism, might participate in the retroaldol elimination of acetyl-CoA performed by msSal via protein–protein interaction through the DUF35 domain of Shy as it is described in *Sphingomonas* sp. strain Ibu-2 and *Thauera aromatica* strain K172 (Leuthner & Heider, 2000; Murdoch & Hay, 2013). In this sense, msShy contains two conserved domains DUF35 (domain of unknown function 35) homologous to those from ChsH2

from *Mycobacterium tuberculosis* that is known to form a functional thiolase complex with Ltp2 (Aggett et al., 2019). However, unlike ChsH2, msShy does not contain the MaoC-like domain suggesting that it might not have the dehydratase activity and could be mainly involved in stabilizing the binding of an unknown hydratase to catalyse the hydration and retroaldol reactions of 3-oxochol-4,22-dien-24-oyl-CoA (Compound 2) (Figure 2). This interaction has been proposed as well in *M. neoaurum* between *aft1* coding an aldolase and *chsH4* coding an acyl dehydratase containing only one DUF354 domain and, as in msShy, the MaoC-like domain is absent (Zhao et al., 2023).

Boosting 4-HBC production in recombinant strains

To reroute the metabolic flux towards the production of 4-HBC avoiding AD accumulation, we overexpressed the aldolase msSal involved in the retroaldol elimination of an acetyl-CoA from the C5-side chain intermediate 22-OH-BCN-CoA. The MS6039-5941-5903 strain was transformed with plasmid pMVmsSal carrying the *MSMEG_6561* gene coding the Sal-like aldolase, under

the control of the *hsp60* promoter. The resulting strain MS6039-5941-5903 (pMVmsSal) showed an impaired growth when cultured in the biotransformation medium in the presence of cholesterol (Figure 6A), causing a lower 4-HBC production than the strain MS6039-5941-5903 (pMV261) carrying the empty vector pMV261 (Figure 6B). This result suggests that increasing Sal levels can generate an excess of the toxic intermediate 3-OPA aldehyde (Figure 2). It is known that the presence of an excess of aldehydes could be detrimental for bacterial cells due to their high reactivity and, therefore, microbial growth is usually inhibited in the presence of these compounds (Kunapur et al., 2014; Kunapur & Prather, 2015; Zaldivar et al., 1999). This result also suggests that the oxidation of 3-OPA into 4-HBC might be a bottleneck in the production of 4-HBC.

In bile acid degrading microorganisms like *S. degradans* sp. strain Chol1, further degradation of the side chain continues with the oxidation of the aldehyde to the corresponding carboxylic acid catalysed by Sad aldehyde dehydrogenase. Homologous of Sad are present in *C. testosteroni* CNB-2 and *P. putida* DOC21 but absent in the well-studied bile acid degrader *R. jostii* RHA1. In *M. neoarum* strain CCTCC AB2019054 this reaction is catalysed by the N-terminal domain of the 4e reductase mnOpccR (Peng et al., 2021). To identify the enzyme responsible to perform this reaction in *M. smegmatis*, the msOpccR and msRed coding genes (*MSMEG_1623* and *MSMEG_6563*, respectively) were cloned into a mycobacterial replicative plasmid forming

an operon structure with the *MSMEG_6561* gene coding msSal. The resulting plasmids were transformed in the MS6039-5941-5903 strain rendering MS6039-5941-5903 (pMVmsSal-msOpccR) and MS6039-5941-5903 (pMVmsSal-msRed) strains.

When both recombinant strains were cultured in the biotransformation medium in the presence of cholesterol as described before, they showed a normal growth comparable to that of the parental strain (Figure 6A), suggesting that the plasmids expressing both reductases can reduce the 3-OPA toxicity.

Regarding the 4-HBC production, the strain MS6039-5941-5903 (pMVmsSal-msRed), achieves a concentration of 0.291 g/L at 96 h, which is slightly higher than the concentration of 0.258 g/L achieved after 96 h by MS6039-5941-5903 (pMV261), i.e., the parental strain carrying the empty vector (Figure 6B). However, the MS6039-5941-5903 (pMVmsSal-msOpccR) strain only accumulated 0.246 g/L (Figure 6B).

Unfortunately, the presence of AD was still detected as a by-product in the biotransformation assays performed with the recombinant strains expressing msRed and/or msOpccR. Figure 6C shows an AD accumulation of 9–19 mg/L after 96 h, which is similar to the AD concentration obtained in the parental strain carrying the empty vector (12 mg/L), suggesting that the expression of msRed or msOpccR reductases separately do not avoid AD accumulation in the 4-HBC-producing strains. However, due to the increase of the msRed coding gene dose, the MS6039-5941-5903 (pMVmsSal-msRed)

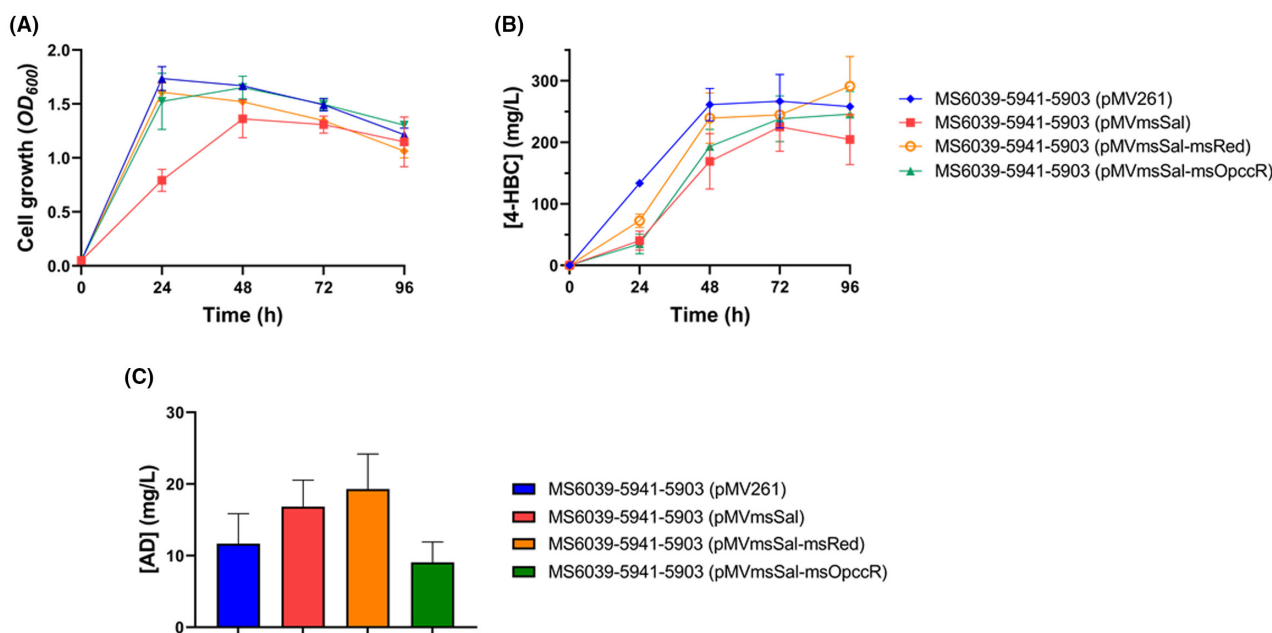


FIGURE 6 (A) Growth curve of MS6039-5941-5903 (pMV261), MS6039-5941-5903 (pMVmsSal), MS6039-5941-5903 (pMVmsSal-msRed) and MS6039-5941-5903 (pMVmsSal-msOpccR) strains in the biotransformation medium. (B) 22-hydroxy-23,24-bisnorchol-4-ene-3-one (4-HBC) production by MS6039-5941-5903 (pMV261), MS6039-5941-5903 (pMVmsSal), MS6039-5941-5903 (pMVmsSal-msRed) and MS6039-5941-5903 (pMVmsSal-msOpccR) strains from 0.4 g/L of cholesterol in shake flasks. (C) 4-androstene-3,17-dione (AD) accumulation by MS6039-5941-5903 (pMV261), MS6039-5941-5903 (pMVmsSal), MS6039-5941-5903 (pMVmsSal-msRed) and MS6039-5941-5903 (pMVmsSal-msOpccR) strains from 0.4 g/L of cholesterol after 96 h cultured in shake flasks.

strain showed an increase of 17% in the 4-HBC molar yield compared to the strain carrying the empty plasmid (Figure 7A). On the other hand, the strain expressing the reductase msOpccR did not show 4-HBC molar yield increase.

Although it has been demonstrated that msOpccR has aldehyde dehydrogenase activity in vitro (Peng et al., 2021), our result suggests that msRed might be the main responsible enzyme for the transformation of 3-OPA to 4-HBC. When msRed is overexpressed together with msSal, the 3-OPA toxicity is reduced and the 4-HBC molar yield is improved, almost reaching 100%, which points to the fact that the overexpression of the aldolase msSal together with the reductase msRed funnels cholesterol into 4-HBC.

Production of 4-HBC in 2L bioreactor

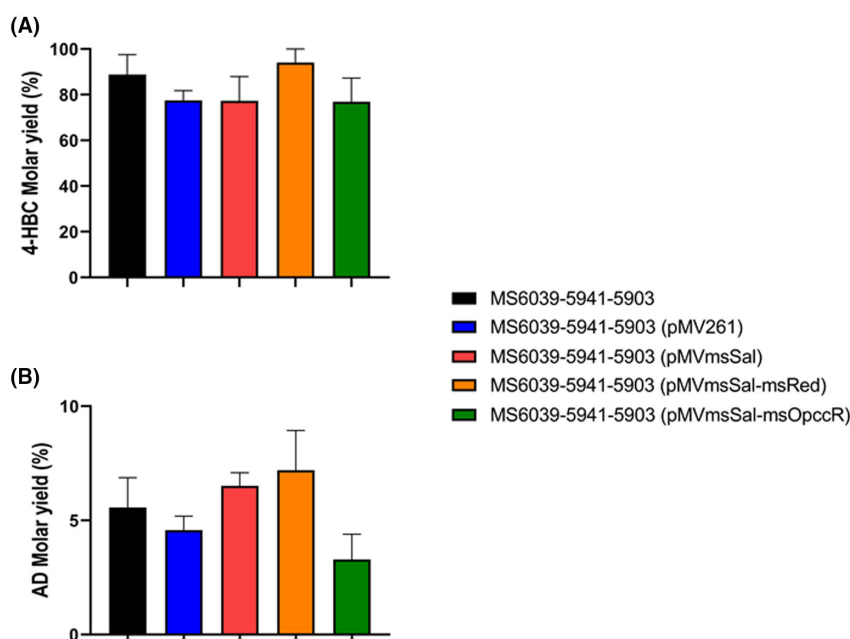
The applicability of the recombinant strains in the production of 4-HBC was evaluated by scaling-up the flask assays to a 2L bioreactor using 14 g/L of phytosterols as substrate. MS6039-5941-5903, MS6039-5941-5903 (pMVmsSal-msRed) and MS6039-5941-5903 (pMVmsSal-msOpccR) strains were cultured for 72 h as described in the Materials and Methods section and samples were taken every 24 h. We observed that MS6039-5941-5903 (pMVmsSal-msRed) displayed a 4-HBC production titer of 12.7 g/L with a productivity of 0.185 g/L h⁻¹ (Figure 8A). This production constitutes an improvement of 127% compared to the MS6039-5941-5903 mutant strain, which renders a 4-HBC production

titer of 5.6 g/L with a productivity of 0.088 g/L h⁻¹ (Figure 8A). Moreover, MS6039-5941-5903 (pMVmsSal-msOpccR) showed also a better performance than the parental strain since it produced 9 g/L of 4-HBC and showed a productivity of 0.140 g/L h⁻¹. This result suggests that msOpccR can perform in vivo the transformation of 3-OPA into 4-HBC as observed in vitro (Peng et al., 2021).

The final AD concentration (2 g/L) was similar at 72 h in the three tested strains although it was slightly higher in the parental strain (Figure 8B). Therefore, the 4-HBC:AD ratio was higher in the strain expressing msSal and msRed (Figure 8C).

In *M. neoaurum* ATCC 25795 the 4-HBC production was enhanced by overexpressing *choM1* and *choM2* encoding cholesterol oxidases, a *cyp125* gene encoding a cytochrome P450 enzyme and a *fadA5* gene encoding a thiolase in a mutant strain lacking KstD, KshA and Hsd4A activities (Xiong et al., 2017). However, high concentrations of AD and 9-OH-AD were still obtained as by-products. In order to reduce these metabolites, the *fadA5* gene was further deleted to prevent the transformation of 22-oxo-BCN-CoA into 3-OPC-CoA, that is an mnOpccR substrate (Figure 2). Although the total content of AD and 9-OH-AD decreased they were not completely removed, because a correlation between the AD and the 4-HBC sub-pathways should exist (Zhao et al., 2023). In *M. fortuitum* ATCC 35855 Hsd4A and Fada5 are found to be the key enzymes in controlling the C19 metabolic pathway of steroids. After knocking out *hsd4A*, the resulting strain, MFKDDhsd4A, accumulated 81.5% 9,22-dihydroxy-23,24-bisnorcholesterol-4-ene-3-one (9-OHBA),

FIGURE 7 (A) Molar yield of 4-HBC production by MS6039-5941-5903, MS6039-5941-5903 (pMV261), MS6039-5941-5903 (pMVmsSal), MS6039-5941-5903 (pMVmsSal-msRed) and MS6039-5941-5903 (pMVmsSal-msOpccR) strains from 0.4 g/L of cholesterol after 96 h cultured in shake flasks. (B) Molar yield of AD production by MS6039-5941-5903, MS6039-5941-5903 (pMV261), MS6039-5941-5903 (pMVmsSal), MS6039-5941-5903 (pMVmsSal-msRed) and MS6039-5941-5903 (pMVmsSal-msOpccR) strains from 0.4 g/L of cholesterol after 96 h cultured in shake flasks.



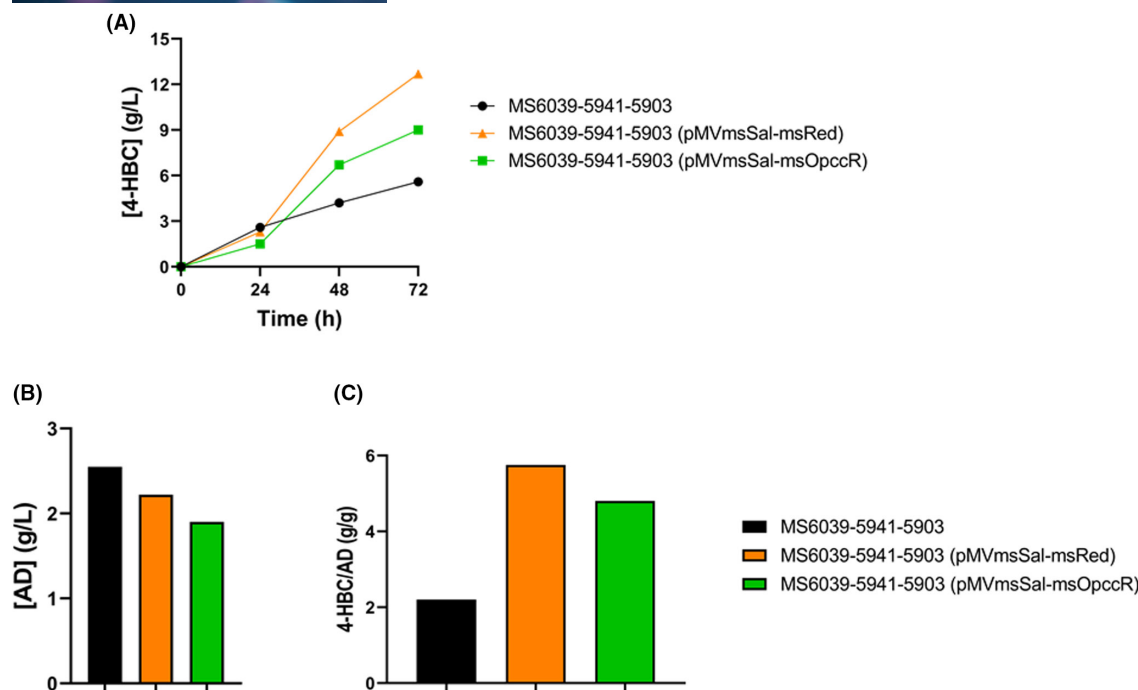


FIGURE 8 (A) 22-hydroxy-23,24-bisnorchol-4-ene-3-one (4-HBC) production over time by MS6039-5941-5903, MS6039-5941-5903 (pMVmsSal-msRed) and MS6039-5941-5903 (pMVmsSal-msOpccR) strains in a bioreactor containing 14 g/L of phytosterols. (B) 4-Androstene-3,17-dione (AD) accumulation by MS6039-5941-5903, MS6039-5941-5903 (pMVmsSal-msRed) and MS6039-5941-5903 (pMVmsSal-msOpccR) strains after 72 h in a bioreactor containing 14 g/L of phytosterols. (C) 22-hydroxy-23,24-bisnorchol-4-ene-3-one (4-HBC):4-Androstene-3,17-dione (AD) ratio after 72 h of transformation in a bioreactor containing 14 g/L of phytosterols.

a 4-HBC hydroxylated derivative, compared with 4.1% 9-OHBA in the strain MFKD while the double mutant MFKDDhsd4ADfadA5 further improved the selectivity of 9-OHBA and decreasing the by-products from 4.19% to 0.90% (Han et al., 2024).

Thus, in order to avoid the accumulation of AD it is necessary to understand its origin in the mutant strains. In β -proteobacteria bile acid degraders the degradation of C3-unit consists in the initiation of a second round of the side-chain degradation and proceeds by the oxidation of the C17-C20 bond by an acyl-CoA dehydrogenase (ACAD). Subsequently, the double bond is hydrated followed by the release of a propionyl-CoA moiety by the action of a thiolase. Several studies have shown that 4-HBC is a dead-end product of sterols in *hsd4A* mutant strains (Liu et al., 2023; Peng et al., 2021; Xu et al., 2016; Zhao et al., 2023). The presence of AD in the biotransformation suggests an oxidative conversion of the aldehyde intermediate 3-OPA to its CoA ester by acyl CoA synthetases turning back the metabolite to the AD pathway, i.e., suffering further dehydrogenation and thiolytic steps after CoA activation that delivers acetyl-CoA and AD. Based on previous data, our hypothesis is that FadA5 would be the most probable enzyme involved in this putative activation although its activity on C3 would be suboptimal since it is mainly acting on substrates containing C7 and C5 side-chains (Nesbitt et al., 2010; Schaefer et al., 2015).

CONCLUSIONS

Our results confirmed that it is possible to improve the performance of the 4-HBC-producing strains not only by deleting the *hsd4A* gene but also by identifying and overexpressing the enzymes that funnel the cholesterol degradation pathway to the 4-HBC end product. This strategy has allowed us to improve a *M. smegmatis* *hsd4A* mutant in order to achieve a higher productivity of 4-HBC, and moreover, it has pointed out new targets for further improvement. The MS6039-5941-5903 (pMVmsSal-msRed) strain displays a high productivity when compared with other 4-HBC producers already described showing a higher 4-HBC:AD ratio, demonstrating that this strain can be useful for industrial applications.

AUTHOR CONTRIBUTIONS

Gabriel Hernández-Fernández: Investigation; writing – original draft; methodology; formal analysis; data curation. **Miguel G. Acedos:** Methodology; investigation; formal analysis; writing – original draft. **Isabel de la Torre:** Investigation; methodology; formal analysis; writing – original draft. **Juan Ibero:** Investigation; writing – original draft. **José L. García:** Conceptualization; funding acquisition; writing – review and editing; supervision. **Beatriz Galán:** Conceptualization; investigation; writing – original draft; writing – review and editing; formal analysis; supervision.

ACKNOWLEDGEMENTS

This work was supported by grants from the Spanish Ministry of Science and Innovation PID2021-125370OB-I00, PID2019-110612RB-I00 from the Ministry of Economy and Competitiveness of Spain and grant from the Community of Madrid and the Structural Funds from European Union (Ref: S2018/BAA-4532 (ALGATEC-CM)). The grant for one of the authors (GHF) with reference FPU21/05101 is gratefully recognized. The authors would like to thank Ana Valencia for the technical assistance and Virginia Rivero and Leonor Rodríguez for their assistance in the Chromatography facility.

CONFLICT OF INTEREST STATEMENT

The authors have no conflict of interest.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available in NCBI at <https://www.ncbi.nlm.nih.gov/>. These data were derived from the following resources available in the public domain: NCBI, <https://www.ncbi.nlm.nih.gov/>.

ORCID

Miguel G. Acedos  <https://orcid.org/0000-0003-4229-5698>

José L. García  <https://orcid.org/0000-0002-9238-2485>

Beatriz Galán  <https://orcid.org/0000-0002-2596-6034>

REFERENCES

- Aggett, R., Mallette, E., Gilbert, S.E., Vachon, M.A., Schroeter, K.L., Kimber, M.S. et al. (2019) The steroid side-chain-cleaving aldolase Ltp2-ChsH2DUF35 is a thiolase superfamily member with a radically repurposed active site. *The Journal of Biological Chemistry*, 294, 11934–11943. Available from: <https://doi.org/10.1074/jbc.RA119.00888>
- Capyk, J.K., Kalscheuer, R., Stewart, G.R., Liu, J., Kwon, H., Zhao, R. et al. (2009) Mycobacterial cytochrome p450 125 (*cyp125*) catalyzes the terminal hydroxylation of c27 steroids. *The Journal of Biological Chemistry*, 284, 35534–35542. Available from: <https://doi.org/10.1074/jbc.M109.072132>
- Donova, M. (2018) Microbiotechnologies for steroid production. *Microbiology Australia*, 39, 126–129. Available from: <https://doi.org/10.1071/MA18039>
- Donova, M. (2021) Microbial steroid production technologies: current trends and prospects. *Microorganisms*, 10, 53–60. Available from: <https://doi.org/10.3390/microorganisms10010053>
- Fabre, M., Mateo, L., Lamaa, D., Baillif, S., Pagès, G., Demange, L. et al. (2022) Recent advances in age-related macular degeneration therapies. *Molecules*, 27, 5089–5153. Available from: <https://doi.org/10.3390/molecules27165089>
- Feller, F.M., Holert, J., Yücel, O. & Philipp, B. (2021) Degradation of bile acids by soil and water bacteria. *Microorganisms*, 9, 1759–1788. Available from: <https://doi.org/10.3390/microorganisms9081759>
- Feng, J., Wu, Q., Zhu, D. & Ma, Y. (2022) Biotransformation enables innovations toward green synthesis of steroidal pharmaceuticals. *ChemSusChem*, 15, e202102399. Available from: <https://doi.org/10.1002/cssc.202102399>

- Fernández-Cabezón, L., Galán, B. & García, J.L. (2018) New insights on steroid biotechnology. *Frontiers in Microbiology*, 9, 1–15. Available from: <https://doi.org/10.3389/fmicb.2018.00958>
- Galán, B., Uhía, I., García-Fernández, E., Martínez, I., Bahillo, E., de la Fuente, J.L. et al. (2017) *Mycobacterium smegmatis* is a suitable cell factory for the production of steroidal synthons. *Microbial Biotechnology*, 10, 138–150. Available from: <https://doi.org/10.1111/1751-7915.12429>
- García, J.L., Uhía, I. & Galán, B. (2012) Catabolism and biotechnological applications of cholesterol degrading bacteria. *Microbial Biotechnology*, 5, 679–699. Available from: <https://doi.org/10.1111/j.1751-7915.2012.00331.x>
- García-Fernández, J., Papavinasasundaram, K., Galán, B., Sassetti, C.M. & García, J.L. (2017) Molecular and functional analysis of the *mce4* operon in *Mycobacterium smegmatis*. *Environmental Microbiology*, 19, 3689–3699. Available from: <https://doi.org/10.1111/1462-2920.13869>
- Han, S., Liu, X., He, B., Zhai, X., Yuan, C., Li, Y. et al. (2024) Efficient production of 9, 22-dihydroxy-23,24-bisnorchol-4-ene-3-one from phytosterols by modifying multiple genes in *Mycobacterium fortuitum*. *International Journal of Molecular Sciences*, 25, 3579. Available from: <https://doi.org/10.3390/ijms25073579>
- Hernández-Fernández, G., Acedos, M.G., García, J.L. & Galán, B. (2023) Identification of the aldolase responsible for the production of 22-hydroxy-23,24-bisnorchol-4-ene-3-one from natural sterols in *Mycobacterium smegmatis*. *Microbial Biotechnology*, 17, e14270. Available from: <https://doi.org/10.1111/1751-7915.14270>
- Holert, J., Jagmann, N. & Philipp, B. (2013) The essential function of genes for a hydratase and an aldehyde dehydrogenase for growth of *pseudomonas* sp. strain Chol1 with the steroid compound cholate indicates an aldolytic reaction step for deacetylation of the side chain. *Journal of Bacteriology*, 195, 3371–3380. Available from: <https://doi.org/10.1128/JB.00410-13>
- Jackson, M., Reinaldo Camacho, L., Gicquel, B. & Guilhot, C. (2001) Gene replacement and transposon delivery using the negative selection marker *sacB*. *Methods in Molecular Medicine*, 54, 59–75. Available from: <https://doi.org/10.1385/1-59259-147-7:059>
- Kunjapur, A.M. & Prather, K.L.J. (2015) Microbial engineering for aldehyde synthesis. *Applied and Environmental Microbiology*, 81, 1892–1901. Available from: <https://doi.org/10.1128/AEM.03319-14>
- Kunjapur, A.M., Tarasova, Y. & Prather, K.L.J. (2014) Synthesis and accumulation of aromatic aldehydes in an engineered strain of *Escherichia coli*. *Journal of the American Chemical Society*, 136, 11644–11654. Available from: <https://doi.org/10.1021/ja506664a>
- Leuthner, B. & Heider, J. (2000) Anaerobic toluene catabolism of *Thauera aromatica*: the *bbs* operon codes for enzymes of beta oxidation of the intermediate benzylsuccinate. *Journal of Bacteriology*, 182, 272–277. Available from: <https://doi.org/10.1128/JB.182.2.272-277.2000>
- Liu, X., He, B., Zhang, J., Yuan, C., Han, S., Du, G. et al. (2023) Phytosterol conversion into C9 non-hydroxylated derivatives through gene regulation in *Mycobacterium fortuitum*. *Applied Microbiology and Biotechnology*, 107, 7635–7646. Available from: <https://doi.org/10.1007/s00253-023-12812-w>
- Murdoch, R.W. & Hay, A.G. (2013) Genetic and chemical characterization of ibuprofen degradation by *Sphingomonas* Ibu-2. *Microbiology*, 159, 621–632. Available from: <https://doi.org/10.1099/mic.0.062273-0>
- Nesbitt, N.M., Yang, X., Fontán, P., Kolesnikova, I., Smith, I., Sampson, N.S. et al. (2010) A thiolase of *Mycobacterium tuberculosis* is required for virulence and production of androstenedione and androstadienedione from cholesterol. *Infection and Immunity*, 78, 275–282. Available from: <https://doi.org/10.1128/IAI.00893-09>

- Nunes, V.O., Vanzellotti, N., Fraga, J.L., Pessoa, F.L.P., Ferreira, T.F. & Amaral, P.F.F. (2022) Biotransformation of phytosterols into androstenedione-a technological prospecting study. *Molecules*, 27, 3164. Available from: <https://doi.org/10.3390/molecules27103164>
- Ouellet, H., Guan, S., Johnston, J.B., Chow, E.D., Kells, P.M., Burlingame, A.L. et al. (2010) *Mycobacterium tuberculosis* CYP125A1, a steroid C27 monooxygenase that detoxifies intracellularly generated cholest-4-en-3-one. *Molecular Microbiology*, 77, 730–742. Available from: <https://doi.org/10.1111/j.1365-2958.2010.07243.x>
- Peng, H., Wang, Y., Jiang, K., Chen, X., Zhang, W., Zhang, Y. et al. (2021) A dual role reductase from phytosterols catabolism enables the efficient production of valuable steroid precursors. *Angewandte Chemie*, 133, 5474–5480. Available from: <https://doi.org/10.1002/ange.202015462>
- Schaefer, C.M., Lu, R., Nesbitt, N.M., Schiebel, J., Sampson, N.S. & Kisker, C. (2015) FadA5 a thiolase from *mycobacterium tuberculosis*: a steroid-binding pocket reveals the potential for drug development against tuberculosis. *Structure*, 1993(23), 21–33. Available from: <https://doi.org/10.1016/j.str.2014.10.010>
- Snapper, S.B., Melton, R.E., Mustafa, S., Kieser, T. & Jacobs, W.R. (1990) Isolation and characterization of efficient plasmid transformation mutants of *Mycobacterium smegmatis*. *Molecular Microbiology*, 4, 1911–1919. Available from: <https://doi.org/10.1111/j.1365-2958.1990.tb02040.x>
- Stover, C.K., de la Cruz, V.F., Fuerst, T.R., Burlein, J.E., Benson, L.A., Bennett, L.T. et al. (1991) New use of BCG for recombinant vaccines. *Nature*, 351, 456–460. Available from: <https://doi.org/10.1038/351456a0>
- Uhía, I., Galán, B., Morales, V. & García, J.L. (2011) Initial step in the catabolism of cholesterol by *Mycobacterium smegmatis* mc2 155. *Environmental Microbiology*, 13, 943–959. Available from: <https://doi.org/10.1111/j.1462-2920.2010.02398.x>
- Wang, X.-X., Ke, X., Liu, Z.-Q. & Zheng, Y.-G. (2022) Rational development of mycobacteria cell factory for advancing the steroid biomanufacturing. *World Journal of Microbiology and Biotechnology*, 38, 191–207. Available from: <https://doi.org/10.1007/s11274-022-03369-3>
- Wang, Z., Qiu, H., Chen, Y., Chen, X., Fu, C. & Yu, L. (2024) Microbial metabolism of diosgenin by a novel isolated *Mycolicibacterium* sp. HK-90: a promising biosynthetic platform to produce 19-carbon and 21-carbon steroids. *Microbial Biotechnology*, 17, e14415. Available from: <https://doi.org/10.1111/1751-7915.14415>
- Wipperman, M.F., Sampson, N.S. & Thomas, S.T. (2014) Pathogen rod rage: cholesterol utilization by *Mycobacterium tuberculosis*. *Critical Reviews in Biochemistry and Molecular Biology*, 49, 269–293. Available from: <https://doi.org/10.3109/10409238.2014.895700>
- Xiong, L.-B., Liu, H.-H., Xu, L.-Q., Sun, W.-J., Wang, F.-Q. & Wei, D.-Z. (2017) Improving the production of 22-hydroxy-23,24-bisnorchol-4-ene-3-one from sterols in *Mycobacterium neoaurum* by increasing cell permeability and modifying multiple genes. *Microbial Cell Factories*, 16, 89–99. Available from: <https://doi.org/10.1186/s12934-017-0705-x>
- Xu, L.-Q., Liu, Y.-J., Yao, K., Liu, H.-H., Tao, X.-Y., Wang, F.-Q. et al. (2016) Unraveling and engineering the production of 23,24-bisnorcholenic steroids in sterol metabolism. *Scientific Reports*, 6, 21928–21931. Available from: <https://doi.org/10.1038/srep21928>
- Yang, B., Zhao, T., Liu, Y. & Lu, B. (2023) Guidelines for extraction and quantitative analysis of phytosterols and oxidation products. *eFood*, 4, e108. Available from: <https://doi.org/10.1002/efd.2108>
- Yücel, O., Borgert, S.R., Poehlein, A., Niermann, K. & Philipp, B. (2019) The 7 α -hydroxysteroid dehydratase Hsh2 is essential for anaerobic degradation of the steroid skeleton of 7 α -hydroxyl bile salts in the novel denitrifying bacterium *Azoarcus* sp. strain Aa7. *Environmental Microbiology*, 21, 800–813. Available from: <https://doi.org/10.1111/1462-2920.14508>
- Zaldivar, J., Martinez, A. & Ingram, L.O. (1999) Effect of selected aldehydes on the growth and fermentation of ethanologenic *Escherichia coli*. *Biotechnology and Bioengineering*, 65, 24–33.
- Zhao, A., Zhang, X., Li, Y., Wang, Z., Lv, Y., Liu, J. et al. (2021) *Mycolicibacterium* cell factory for the production of steroid-based drug intermediates. *Biotechnology Advances*, 53, 107860–107889. Available from: <https://doi.org/10.1016/j.biotechadv.2021.107860>
- Zhao, Y.-Q., Liu, Y.-J., Song, L., Yu, D., Liu, K., Liu, K. et al. (2023) Unravelling and engineering an operon involved in the side-chain degradation of sterols in *Mycolicibacterium neoaurum* for the production of steroid synthons. *Biotechnology for Biofuels and Bioproducts*, 16, 121–138. Available from: <https://doi.org/10.1186/s13068-023-02376-2>

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Hernández-Fernández, G., Acedos, M.G., de la Torre, I., Ibero, J., García, J.L. & Galán, B. (2024) Improving the production of 22-hydroxy-23,24-bisnorchol-4-ene-3-one in *Mycolicibacterium smegmatis*. *Microbial Biotechnology*, 17, e14551. Available from: <https://doi.org/10.1111/1751-7915.14551>

ANEXO III. Publicación 3

Unraveling the key drivers of bacterial progesterone degradation

Gabriel Hernández-Fernández,^{1,2} Juan Ibero,¹ José L. García,¹ Beatriz Galán¹

AUTHOR AFFILIATIONS See affiliation list on p. 19.

ABSTRACT Progesterone (PROG) is one of the most ubiquitous sexual hormones found as a pollutant in soil and water systems. Despite the fact that PROG can be degraded by various bacterial species, the pathways leading to its complete oxic mineralization remain unknown. In this study, we investigated bacterial progesterone catabolism using the steroid-degrading bacterium *Caenibius tardaugens* as a model, in which we demonstrated its capacity to degrade progestogens. The transcriptomic analyses in the presence of PROG showed the overexpression of *EGO55_13845* and *EGO55_13860* genes, coding a Baeyer-Villiger monooxygenase (BVMO) and a luciferase-like monooxygenase (LLM), respectively. Both genes are located next to two regulatory proteins forming a small gene cluster (named *pdc*) that can be found in other steroid-degrading bacteria. Mutagenic analyses and gene complementation allowed ascertaining that *EGO55_13845* and *EGO55_13860* genes are involved in PROG degradation. To assess their enzymatic activities, both proteins were overexpressed in *Escherichia coli*, showing that they catalyze a Baeyer-Villiger monooxygenation of PROG, resulting in the production of testosterone acetate. They are also active on 1,2-dehydroprogesterone, an intermediate in PROG degradation, converting it into boldenone acetate. BVMO and LLM enzymes are functionally redundant, as each can replace the other in metabolizing PROG. The presence of two functionally redundant BVMO and LLM enzymes in *C. tardaugens* can be explained by their distinct substrate preferences. Our results settle for the first time the genetic and biochemical basis for explaining how PROG is recognized and channeled into the 9,10-seco degradation pathway in a PROG-degrading bacterium.

IMPORTANCE This study investigates for the first time the key steps in bacterial progesterone (PROG) degradation, revealing new insights into the process. The main stages of PROG degradation were examined in the bacterium *Caenibius tardaugens*. The conducted transcriptomic analysis allows identifying the progesterone degradation cluster *pdc*, which is also present in other related bacteria. We demonstrated that Baeyer-Villiger monooxygenase and luciferase-like monooxygenase encoded within the *pdc* cluster catalyze the PROG Baeyer-Villiger monooxygenation, producing testosterone acetate. The activity redundancy can be explained by the difference in substrate specificity of each enzyme.

KEYWORDS progesterone, endocrine disruptors, Baeyer-Villiger monooxygenases, luciferase-like monooxygenase, testosterone acetate

Endocrine disruptors (EDCs) are natural and synthetic compounds that interfere with endocrine systems, causing adverse effects in humans and leading to the feminization or masculinization of freshwater wildlife exposed to polluted rivers (1–3). Among EDCs, natural estrogens (e.g., 17 β -estradiol), androgens (e.g., testosterone [TES]), and progestogens (e.g., progesterone [PROG]) are the most widespread sex hormones found

Editor Caroline S. Harwood, University of Washington School of Medicine, Seattle, Washington, USA

Address correspondence to Beatriz Galán, bgalan@cib.csic.es.

The authors declare no conflict of interest.

See the funding table on p. 19.

Received 7 April 2025

Accepted 6 May 2025

Published 30 May 2025

Copyright © 2025 Hernández-Fernández et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

as pollutants in soil and water. These contaminants enter the environment through human, domestic, and farm animal excretions, as well as pharmaceutical waste (4, 5).

PROG is an endogenous 21-carbon steroid hormone synthesized from cholesterol in the corpus luteum of the ovaries and the placenta during pregnancy. It plays a crucial role in the female menstrual cycle, pregnancy, and embryogenesis in humans and other species (6). Due to its functional role, PROG is extensively used in oral contraceptives alongside the synthetic estrogen 17 α -ethynylestradiol, as well as in hormone treatments to prevent miscarriage, hormone replacement therapy for menopausal women, and veterinary medicine (6). As a result, PROG is widely consumed, excreted through urine by humans, livestock, and other vertebrates, and subsequently released into the environment. Progestogens have been detected in wastewater treatment plants at concentrations ranging from 0.11 to 110 ng/L, and in surface water at levels between 0.82 and 1.1 ng/L. Their endocrine-disrupting effects on aquatic species have been extensively documented (7). Although progesterone undergoes biodegradation, it can still be detected in the effluents of both industrial and municipal wastewater and even in the receiving rivers (8, 9). In this sense, it has been documented that minor transformation of progestogens can create other classes of steroids with different biological activities like potent androgens (10–12).

Gut microbiota can convert glucocorticoids into progestins, via 21-dehydroxylation, in the presence of H₂ gas (13). This is an example of cooperative metabolism dependent on the presence of certain bacteria such as *Eggerthella lenta* and *Gordonibacter pamelaceae*, which carry out the 21-dehydroxylation, and *Escherichia coli*, which provides H₂. More specifically, PROG can be metabolized in the gut to 5 α - and 5 β -pregnanolone, 20-dihydroxyprogesterone, 11-desoxycorticosterone, and 17 α -hydroxyprogesterone, compounds that have different targets and/or functions in the human body (14).

Microbial degradation can contribute to eliminating natural or synthetic steroid hormones from polluted systems, and the persistence and fate of some of these compounds have been studied previously (15, 16). Nevertheless, only a limited number of organisms can mineralize steroids (17–19). The pathways for the complete oxic mineralization of TES and 17 β -estradiol (E2) have been studied in detail in several bacteria (20–24). However, although bacteria capable of degrading PROG have been described (25–27), there is limited information on the genes involved in the first steps of its mineralization, and only a few biochemical studies have been published. Liu et al. studied the aerobic degradation of PROG and norgestrel, detecting 4-androstene-3,17-dione (AD), androsta-1,4-diene-3,17-dione (ADD), 17 β -boldenone, testosterone, 6,7-dehydroprogesterone, and 1,2-dehydroprogesterone (1,2-dPROG) as intermediates in the culture medium (28). They proposed that PROG undergoes different transformations ending up in ADD or 17 β -boldenone, which are finally mineralized. Horinouchi et al. suggested that *Comamonas testosteroni* degrades PROG through the well-known 9,10-seco pathway used as well in the bacterial mineralization of sterols, bile acids, and androgens, but the genes and the enzymes that channel PROG into this degradative pathway were not described (20).

It is known that some fungal species can degrade and cleave the side chain of PROG, transforming it into potent androgens like testololactone, which has anticancer activity, through several steps mediated by Baeyer-Villiger monooxygenases (BVMOs). However, the activity has been demonstrated only through whole cell biotransformations, and the enzymes responsible for these biochemical steps have not been isolated and identified so far (29, 30). Based on the identification of the biotransformation intermediates, it has been postulated that during side chain degradation of PROG, oxygen is typically introduced between C17 and C20 to form 17-acetate, followed by cleavage of the side chain, leading to the formation of androgens with a C19 skeleton, particularly TES. TES is then oxidized to form AD, a C19 compound with a 3-keto group. The final and critical step is the conversion of AD to testololactone. In this sense, BVMOs catalyze the insertion of an oxygen atom into the 3-keto group of AD, which results in the formation of a C19 lactone ring in testololactone (29).

The activity of a bacterial steroid monooxygenase, mediated by BVMOs, introducing an oxygen atom between the C17 and C20 carbons of PROG, was first demonstrated in the gram-positive actinomycete *Rhodococcus rhodochrous* (formerly *Nocardia corallina*) by Miyamoto et al. (31). The gene encoding this enzyme was cloned by Morii et al. (32), and the protein was structurally characterized by Franceschini et al. (33). The properties of the enzyme differed considerably from those of the corresponding fungal monooxygenases because it catalyzed oxidative esterification of PROG to testosterone acetate (TES-Ac) as well as its 11(α / β)-hydroxy derivatives, but the lactonization of AD was not observed. Although these studies confirm the presence of a PROG-transforming enzyme in this microorganism, they did not describe the PROG catabolic pathway. Other BVMOs from *Xanthobacter*, *Comamonas*, and *Rhodococcus* have been cloned and assayed by Van Beilen et al. (34), but none of them showed activity on PROG.

In this work, we have studied the PROG catabolism using as a model the gram-negative alphaproteobacterium *Novosphingobium tardaugens* NBRC 16725 (described as strain ARI-1) and now renamed as *Caenibius tardaugens* (35). This bacterium was isolated at a sewage treatment plant in Tokyo, Japan, due to its ability to use estrogens as the only carbon and energy source (36). Its curated genome is now available (37), and it was used as a model to elucidate the E2 and TES degradation pathways (21, 22) (Fig. 1). Here, for the first time, we identify the genes and characterize the enzymes involved in the PROG degradation pathway in a bacterium capable of degrading a large variety of endocrine disruptors.

RESULTS

C. tardaugens grows on a wide range of steroids

Figure 2 shows that *C. tardaugens* is able to use different steroids as sole carbon and energy sources. *C. tardaugens* efficiently uses different C18 and C19 steroids without side chains such as E2, estrone, and estriol (22), AD, ADD, TES, dehydroepiandrosterone (DHEA), TES-Ac, and testosterone cypionate (TES-Cyp) (Fig. 2A), as well as steroids with different side chain lengths such as deoxycholate (dCholate), cholate, pregnenolone (PREG), and PROG (Fig. 2B). Cyclodextrins (CDX) used to solubilize steroids are not metabolized by this strain. These results confirm that *C. tardaugens* is the only tested gram-negative bacterium capable of degrading a wide variety of steroid substrates. This ability suggests that it is equipped with a great arsenal of metabolic pathways and makes it an excellent model for studies on the biodegradation of steroidal endocrine disruptors. To date, we have identified and characterized the metabolic pathways to degrade estrogens (C18 steroids) and androgens (C19 steroids) in this strain (21, 22), but the key drivers responsible for degrading steroids containing side chains at C17, such as progesterone, have not been described yet. Based on this, we decided to investigate the pathway responsible for PROG metabolism. Although PROG is one of the central EDC pollutants, to date, the microbial genes involved in its degradation have not been identified and/or described.

A transcriptomic analysis of *C. tardaugens* grown on progesterone reveals the presence of the *pdc* cluster

To identify the genes involved in the degradation of PROG, we performed RNA-seq analysis. Total RNA was extracted from *C. tardaugens* cultured in PROG and in TES, as the control condition. The differential expression analysis of these data yielded 36 differentially expressed genes (DEGs) out of the 4,080 genes encoded in the genome (Table S1). Among the 36 DEGs, 33 were upregulated with a fold change (FC) > 2 and only 3 were downregulated in PROG with FC < -2 (Fig. S1), suggesting that there are few differences between the TES and PROG degradation pathways, most probably because both might converge at a certain point. Previous transcriptomic data of this strain allowed us to describe in detail the TES degradation cluster (*SD* cluster) and the E2 degradation cluster (*edc* cluster) of *C. tardaugens* through differential expression analysis in the presence of TES or E2 vs pyruvate (20, 21).

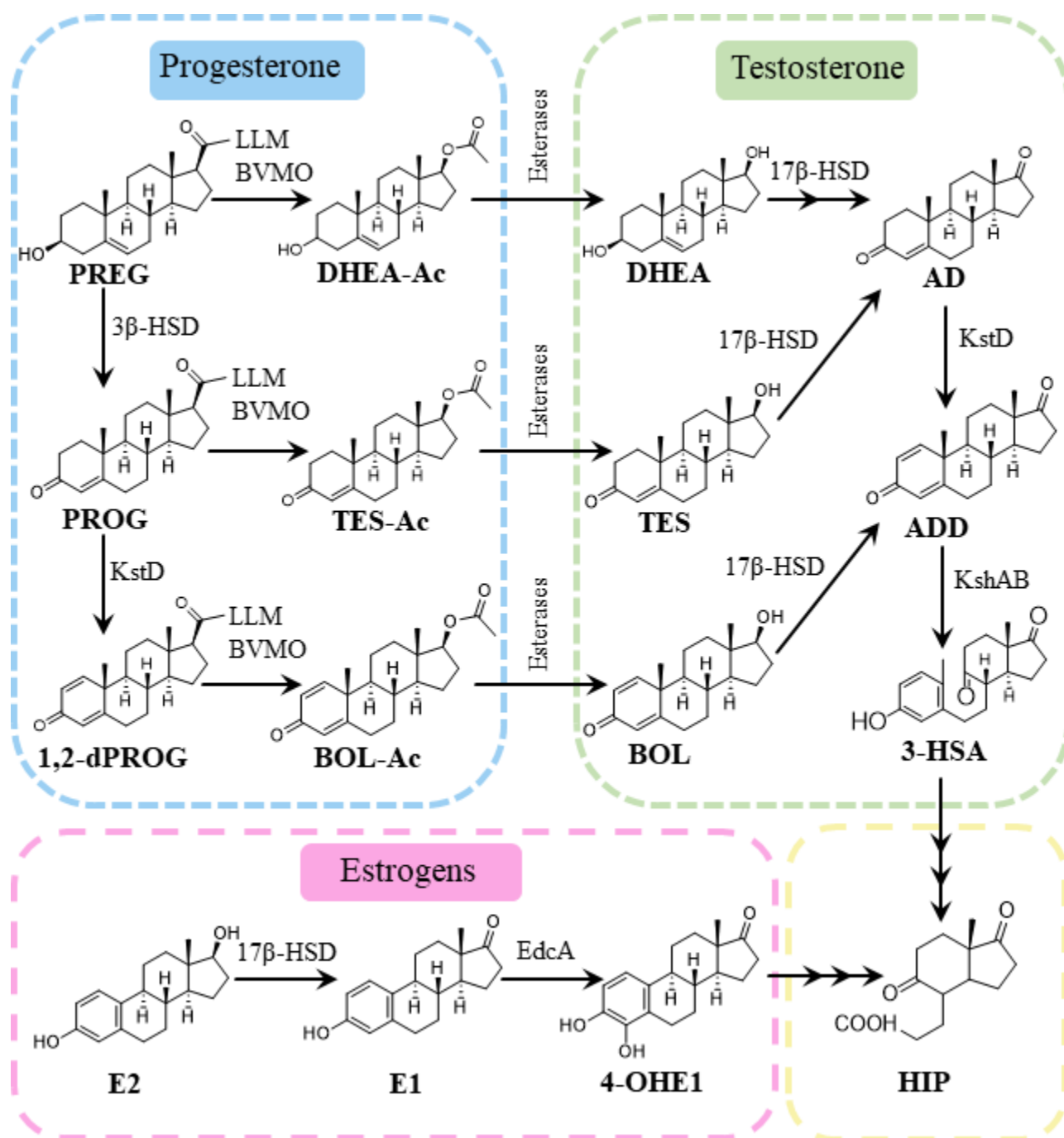


FIG 1 Scheme of all the steroid-degrading routes described in *C. talarum*. The progesterone degradation pathway described in this work is highlighted in blue. The testosterone degradation pathway is marked in green while the estrogen degradation pathway is colored in pink, and both are described by Ibero et al. (20, 21). PREG, pregnenolone; DHEA-Ac, androstenediol acetate; BOL-Ac, boldenone acetate; DHEA, dehydroepiandrosterone; BOL, boldenone; 3-HSA, 3-hydroxy-9,10-secoandrost-1,3,5(10)-triene-9,17-dione; E1, estrone; 4-OHE1, 4-hydroxyestrone; HIP, 3α-H-4α(3'-propanoate)-7α-β-methylhexahydro-1,5-indanedione; LLM, luciferase-like monooxygenase; 3β-HSD, 3-hydroxysteroid dehydrogenase; 17β-Hsd, 17β-hydroxysteroid dehydrogenase; KstD, 3-ketosteroid-Δ1-dehydrogenase; KshAB, 3-ketosteroid 9α-hydroxylase; EdcA, estrone 4-hydroxylase.

Figure 3 presents a comparison of the transcripts per kilobase million (TPMs) of each gene expressed in PROG vs TES, highlighting the most upregulated genes that may play a role in PROG metabolism. The highest level of upregulation FC > 5 was observed in seven genes grouped in three different clusters: *EGO55_04225-04235*,

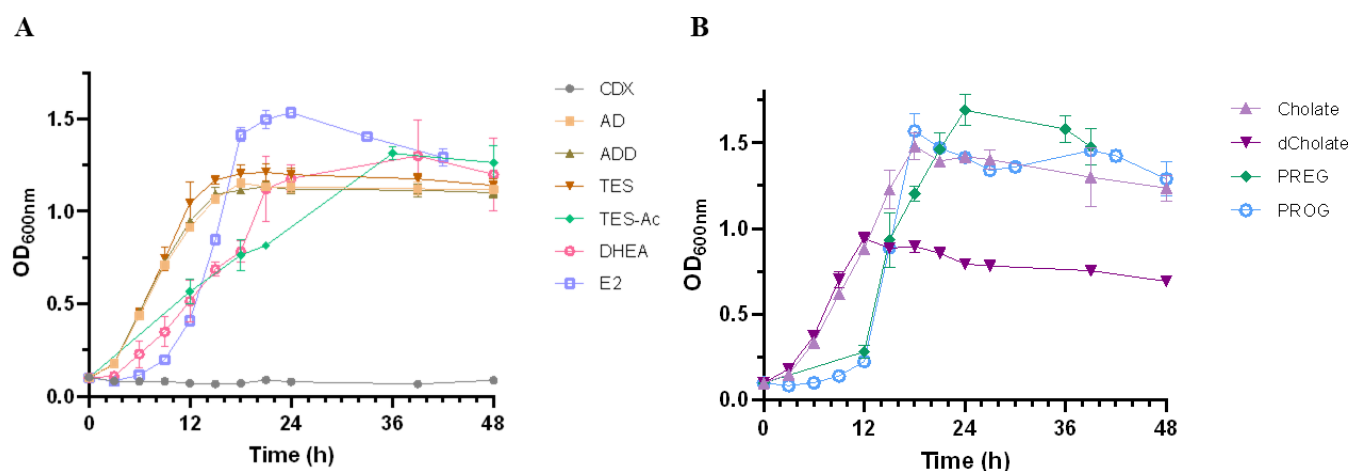


FIG 2 (A) Growth curves of *C. taradagens* NBRC 16725 on different C19 steroids. The values represented are the mean of three independent biological replicates. (B) Growth curves of *C. taradagens* NBRC 16725 on different steroids with side chain. Cholate, sodium cholate. The values represented are the mean of three independent biological replicates.

EGO55_13845–*EGO55_13860*, and *EGO55_14520*–*EGO55_14525*. The *EGO55_04225*–*04235* cluster encodes a hypothetical protein, a Dabb family protein, and a NADP-dependent oxidoreductase, showing an FC of 9, 15, and 11, respectively (Table S1). *EGO55_14520* encodes a tyrosine phosphatase with an FC of 8, while *EGO55_14525* encodes an short-chain dehydrogenase/reductase (SDR) family oxidoreductase with an FC of 15 (Table S1). Notably, the highest FC values were found for the cluster containing the *EGO55_13845* and *EGO55_13860* genes, encoding a BVMO and a luciferase-like monooxygenase (LLM) with values of 139 and 32, respectively. Within this cluster, two additional

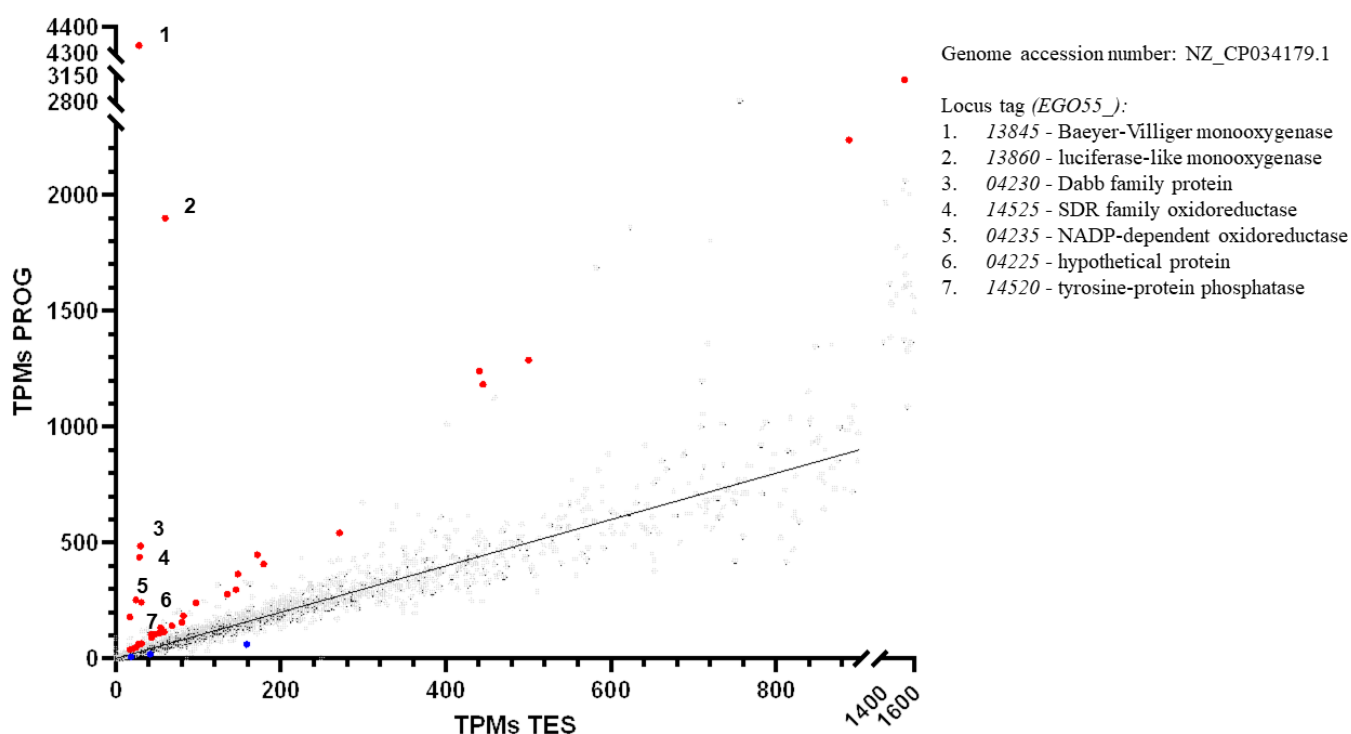


FIG 3 Distribution of the DEGs according to their expression in progesterone or testosterone expressed in TPMs. Red dots represent DEGs with FC > 2 and false discovery rate (FDR) < 0.05, while blue dots represent DEGs with FC < -2 and FDR < 0.05. Gray dots correspond to genes showing 2 > FC > -2. The seven DEGs more overexpressed are marked from 1 to 7 with their locus tag. SDR: short-chain dehydrogenase/reductase.

genes, *EGO55_13850* and *EGO55_13855*, encoding transcriptional regulators, are not induced in PROG, displaying FC values of -1 and 3 , respectively (Table S1). This cluster is 12.8 kb distant from the *SD* cluster (TES degradation) (20) and 51.4 kb distant from the *edc* cluster (E2 degradation) (21), and therefore, in principle, the detected cluster cannot be directly associated with these degradation pathways. Considering the high level of upregulation of the *EGO55_13845*–*EGO55_13860* cluster, we have named the *pdc* cluster that stands for *PROG* *degradation* *cluster* and selected it for further studies.

Annotation of the *pdc* cluster

The *pdc* cluster of *C. tarda* spans 4.6 kb and comprises four genes (Fig. 4). *EGO55_13845* encodes a BVMO that contains a transmembrane domain and a non-cytoplasmic domain consisting of a NAD(P)-binding Rossmann-like domain and an FAD-binding domain. Baeyer-Villiger monooxygenases catalyze the oxidation of a ketone group to an ester group using NAD(P)H and molecular oxygen, with the release of a water molecule (38). BVMOs can be classified into two main groups depending on whether they bind FAD or FMN as cofactors (38, 39). Type I BVMOs have two Rossman folds enabling the protein to bind FAD and NAD(P)H, so that a single protein can perform the oxidation and reduction activities (33, 40). On the other hand, type II BVMOs have the oxygenating and the reducing components in two different proteins: the reducing component binds NADH and reduces FMN to FMNH₂, and the oxygenating component contains an eight-stranded α/β -barrel to bind FMNH₂ and molecular oxygen to perform the oxidation (41). The Baeyer-Villiger reaction has been described in several BVMOs from fungi and other bacteria due to their interest as a high regio-, stereo-, and enantioselect-

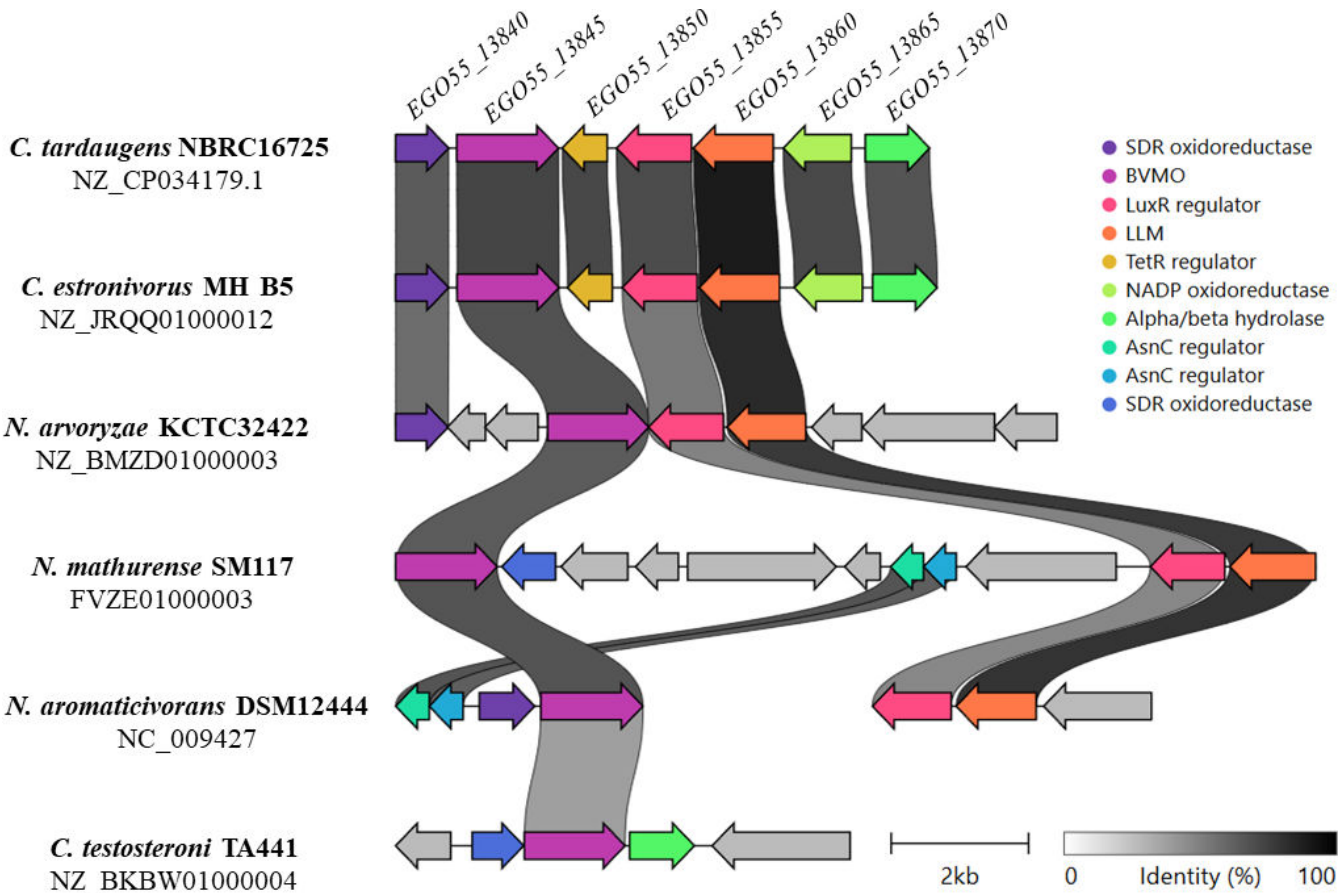


FIG 4 Gene organization of the *pdc* cluster in *C. tarda* compared to other related bacteria.

tive biocatalyst over different aliphatic and aryl ketones and steroids. BVMOs catalyze the insertion of an oxygen atom into ketone substrates to generate lactones or esters.

EGO55_13845 encodes a type I BVMO with a 30.4% amino acid sequence identity to that of *R. rhodochrous* that transforms PROG into TES-Ac *in vitro* (31–34). A large number of putative type I BVMOs are encoded in the *C. tardaogens* genome, sharing varying degrees of similarity. One of the most intriguing is the BVMO encoded by the *EGO55_06865* gene, located 1.4 Mbp away from the *pdg* cluster. This BVMO shares a 47.6% amino acid sequence identity with the BVMO from *R. rhodochrous* and a 53.0% amino acid sequence identity with the BVMO encoded by *EGO55_13845*. However, the transcriptomic analysis showed that the *EGO55_06865* gene is not differentially expressed. Moreover, the TPMs for *EGO55_06865* were around 50, while those for *EGO55_13845* were above 4,000 (Table S1), suggesting that *EGO55_06865* is likely not involved in PROG metabolism, or at least cannot be induced under the culture conditions used in this study. BVMOs highly similar to *EGO55_13845* can be found in other well-known steroid-degrading bacteria, such as *Croceicoccus estronivorus* MH-B5, *Novosphingobium aromaticivorans* DSM 12444, and *C. testosteroni* TA441, sharing 79%, 71%, and 47% protein sequence identity, respectively (Fig. 4). Interestingly, the BVMO (CTTA_RS14345) from *C. testosteroni* TA441 shares 47% identity with *EGO55_13845* but a 55% identity with *EGO55_06865*. Two putative regulatory genes are located in the *pdg* cluster. *EGO55_13850* gene encodes a TetR/AcrR family transcriptional regulator, whereas the *EGO55_13855* gene encodes a helix-turn-helix transcriptional regulator belonging to the LuxR family. The *EGO55_13850* gene is slightly downregulated, showing an FC of -1 (Table S1). This might suggest that the encoded regulator may operate as a repressor of the *pdg* cluster, according to the main function of the TetR-like regulator (42). In contrast, the expression of the *EGO55_13855* gene was upregulated in the presence of PROG with an FC of 3 (Table S1), suggesting that the encoded regulator might act as an activator. Curiously, the TES degradation pathway in *C. testosteroni* TA441 is induced by TesR, a similar LuxR-type transcriptional regulator (43). Both regulators have homologous proteins in other bacteria with an amino acid sequence identity ranging from 55 to 78% (Fig. 4). However, no homologous gene to *EGO55_13850* was found in the genome of *Novosphingobium arvoryzae* KCTC 32422 having a similar *pdg* cluster to that of *C. tardaogens* (Fig. 4).

Finally, the *EGO55_13860* gene encodes an LLM containing one luciferase-like domain characterized by displaying a triose-phosphate isomerase (TIM)-barrel folding, which is the typical domain present in type II BVMOs (41). There are only a few biochemical studies involving this kind of BVMOs, probably because the catalysis involves a two-protein complex, one monooxygenase and one reductase. The LLM *EGO55_13860* present in the *pdg* cluster might be the oxygenation component in the Baeyer-Villiger reaction requiring a flavin reductase. The *EGO55_13865* gene located next to *EGO55_13860* encoded a protein annotated as a NADP-dependent oxidoreductase, and according to the transcriptomic analysis, its expression remains below 30 TPMs in the presence of PROG with an FC around -1 . Therefore, we cannot ascribe this reductase to the second component providing the flavin reductase required for the LLM activity. There is only one flavin reductase within the *C. tardaogens* genome encoded by the *EGO55_18005* gene located distant from the *pdg* cluster. Its expression in the transcriptome is around 100 and 170 TPMs in PROG and TES conditions, respectively, with an FC of -1 (Table S1). These values are far from the TPMs $> 1,700$ and FC > 30 observed for *EGO55_13860*, but the low expression levels might be enough for its reduction activity, avoiding excessive metabolic burden in the bacterial cell. The LLM *EGO55_13860* has no other homologous proteins in the *C. tardaogens* genome. However, LLM homologs can be found in other steroid-degrading bacteria, such as MB02_15840, IE113_RS08665, and Saro_3594 from *C. estronivorus* MH-B5, *N. arvoryzae* KCTC 32422, and *N. aromaticivorans* DSM 12444, sharing 92%, 87%, and 88% of protein sequence identity, respectively (Fig. 4). No LLM *EGO55_13860* homologous proteins were found in the PROG-degrading bacterium *C. testosteroni* TA441 (Fig. 4). However, it is worth mentioning

that *Mycobacterium smegmatis* is able to use C19 steroids when the ADD+ cluster is derepressed (44). Interestingly, this ADD+ cluster contains an LLM protein encoded by *MSMEG_4862*, having 28% aa identity with LLM EGO55_13860, which can explain why *M. smegmatis* can use PROG as substrate after activation of the ADD+ cluster (personal communication).

BVMO and LLM are key enzymes in PROG degradation

The *EGO55_13845* and *EGO55_13860* genes encoding the BVMO and LLM enzymes, respectively, were deleted in *C. talarum* in order to prove their role in the PROG degradation pathway. Through bacterial conjugation and using the suicidal plasmid pk18msg (Table 1), harboring the 700 bp upstream and downstream regions of *EGO55_13845* and *EGO55_13860*, two new strains were obtained: *C. talarum* Δ BVMO and *C. talarum* Δ LLM. Both mutants showed a similar growth to that of the wild-type (WT) strain when growing in nutrient broth (NB), suggesting that *EGO55_13845* and *EGO55_13860* are not essential genes for *C. talarum* basal metabolism. However, unexpectedly, when both strains were cultured in the presence of 1.71 mM PROG as sole carbon and energy source, both mutants were able to grow, although a longer lag phase was observed in the *C. talarum* Δ LLM mutant (Fig. 5A). These results might suggest that both enzymes are not involved in the metabolism of PROG, but considering that both enzymes could display similar activities, they can provide an enzymatic redundancy. Even more, considering that BVMO is very similar to that EGO55_06865, this enzyme might take its function.

To determine a possible redundancy in the function of both enzymes, we constructed a double mutant performing the mutation of BVMO in the *C. talarum* Δ LLM strain, yielding the *C. talarum* Δ LLM-BVMO strain. Supporting the redundancy hypothesis, the double mutant did not grow when cultured with 1.71 mM PROG as sole carbon and energy source (Fig. 5A). To further confirm that BVMO and LLM are required to metabolize PROG, the *C. talarum* Δ LLM-BVMO strain was transformed with plasmid pSEVA23_BVMO-LLM, harboring the *EGO55_13845* and *EGO55_13860* genes under the expression of *P_{lexA}* promoter, thus resulting in the new strain *C. talarum* Δ LLM-BVMO (pSEVA23_BVMO-LLM). As control, *C. talarum* Δ LLM-BVMO mutant was transformed with the pSEVA237plexA empty vector, yielding the *C. talarum* Δ LLM-BVMO (pSEVA237plexA) strain. When both strains were cultured with PROG as the sole carbon and energy source, only the complemented *C. talarum* Δ LLM-BVMO (pSEVA23_BVMO-LLM) strain recovered the ability to grow (Fig. 5A), demonstrating that the double mutant has no additional mutations that can impair its growth in PROG.

Using the *C. talarum* Δ LLM and *C. talarum* Δ BVMO mutants, we searched for the presence of intermediate compounds from PROG degradation in the extracellular medium from their cultures in the presence of 1.71 mM PROG. Samples were taken every 3 h, and total steroids were extracted with chloroform to be analyzed by high-performance liquid chromatograph with mass spectrometry (HPLC-MS). After 24 h of culture, PROG is still detected in the culture medium of *C. talarum* Δ LLM mutant when compared to the wild-type strain, but a new peak was detected in the chromatograms (Fig. 5B). The fragmentation pattern of this peak indicated a molecular weight of 313 [M+H]⁺, which might correspond to a dehydrogenated version of PROG. This finding was confirmed by comparing this peak with the peak of commercial 1,2-dPROG (Fig. 5C). In cultures of *C. talarum* Δ LLM, most of the PROG was transformed into 1,2-dPROG after 30 h and no substrate was observed after 39 h (Fig. S2). 1,2-dPROG was also detected after 18 h in the supernatants of cultures of the *C. talarum* Δ BVMO strain, even though this mutant did not show a longer lag phase when compared to the wild-type strain. However, in this case, all 1,2-dPROG and PROG were consumed after 24 h of cultivation (Fig. S3).

In the case of the *C. talarum* Δ LLM-BVMO strain, cells were cultured in a medium containing glucose as carbon and energy source and 1.71 mM PROG to detect possible intermediates. As expected, 1,2-dPROG was detected in the medium, and it remained

TABLE 1 Bacterial strains, plasmids, and primers used in this study

Strain, plasmid, or primer	Genotype and/or description	Reference or sequence
Strains		
<i>C. taylorii</i>		
NBRC 16725	Wild-type strain	(27)
ΔBVMO	NBRC 16725 Δ <i>EGO55_13845</i>	This work
ΔLLM	NBRC 16725 Δ <i>EGO55_13860</i>	This work
ΔLLM-BVMO	NBRC 16725 Δ <i>EGO55_13845–13860</i>	This work
ΔLLM-BVMO	NBRC 16725 Δ <i>EGO55_13845–13860</i> with pSEVA23pLexA- <i>EGO55_13845–13860</i>	This work
(pSEVA23_BVMO-LLM)		
ΔLLM-BVMO	NBRC 16725 Δ <i>EGO55_13845–13860</i> with pSEVA23pLexA	This work
<i>E. coli</i>		
HB101	<i>supE44, ara14, galK2, leuB, lacY1, Δ(gpt-proA)62, rpsL20, xyl-5, mtl-1, recA13, Δ(mcrC-mrr), hsdS20, (rB[−], mB[−]), Sm^R</i>	(45)
DH10B	<i>F[−], mcrA, Δ(mrr, hsdRMS-mcrBC), φ80dlacZΔM15, ΔlacX74, deoR, recA1, araD139, Δ(ara-leu)7697, galU, galK, λ[−], rpsL, endA1, nupG</i>	Invitrogen, UK
DH10B	DH10B with pK18 <i>EGO55_13845</i>	This work
(pK18 <i>EGO55_13845</i>)		
DH10B	DH10B with pK18 <i>EGO55_13860</i>	This work
(pK18 <i>EGO55_13860</i>)		
DH10B (pSEVA23_BVMO-LLM)	DH10B with pSEVA23pLexA- <i>EGO55_13845–13860</i>	This work
DH10B (pSEVA23pLexA)	DH10B pSEVA23pLexA	This work
BL21(DE3)	<i>F[−], ompT, hsdSB (rB[−] mB[−]) gal, dcm λDE3</i> (harboring gene 1 of the RNA polymerase from phage T7 under the <i>P_{lacUV5}</i> promoter)	(46)
BL21(DE3)	BL21(DE3) with pET-29(+) <i>EGO55_13845</i>	This work
(pET-29(+)/BVMO)		
BL21(DE3) (pET-29(+)/LLM)	BL21(DE3) with pET-29(+) <i>EGO55_13860</i>	This work
BL21(DE3) [pET-29a(+)]	BL21(DE3) with pET-29(+)	This work
Plasmids		
pRK600	<i>Cm^R, ColE1 oriV, RP4 oriT</i> ; helper plasmid necessary for triparental matings	(47)
pK18msg	<i>Km^R, ColE oriV, Mob⁺, lacZa, sacB</i> ; vector for gene deletion via homologous recombination	(46)
pK18 <i>EGO55_13845</i>	pK18mobsacB derivative containing the upstream and downstream fragments of <i>EGO55_13845</i> cloned with <i>Bsal</i> for directed mutagenesis	This work
pK18 <i>EGO55_13860</i>	pK18mobsacB derivative containing the upstream and downstream fragments of <i>EGO55_13860</i> cloned with <i>Bsal</i> for directed mutagenesis	This work
pSEVA23pLexA	<i>Km^R, pBBR1 oriV</i> , constitutive expression of <i>gfp</i> gene under the control of the <i>P_{lexA}</i> promoter	(48)
pSEVA23_BVMO-LLM	pSEVA23pLexA where the <i>gfp</i> gene was replaced by <i>EGO55_13845–13860</i> between <i>XbaI</i> and <i>SpeI</i>	This work
pET-29a(+)	Cloning and expression vector, <i>Km^R, ColE1 oriV</i> , T7 promoter	Novagen, Germany
pET-BVMO	pET-29 containing the <i>EGO55_13845</i> gene	This work
pET-LLM	pET-29 containing the <i>EGO55_13860</i> gene	This work
Primers		
<i>EGO55_13845</i> UpFw_	Amplification of the fragment upstream <i>EGO55_13845</i> (<i>Bsal</i>)	5'-CACATGGTCTCTTCTGTGCATATACCGATCA-3'
<i>EGO55_13845</i> UpRv_		5'-CACATGGTCTCGTCATCCCCGTTACTGTTCTGAT-3'
<i>EGO55_13845</i> DwFw_	Amplification of the fragment downstream <i>EGO55_13845</i> (<i>Bsal</i>)	5'-CACATGGTCTCGATGACTTGCTGATATCTGTTGA-3'
<i>EGO55_13845</i> DwRv_		5'-CACATGGTCTCGTTCCCACTATGACACGAAAGC-3'
<i>EGO55_13860</i> UpFw_	Amplification of the fragment upstream <i>EGO55_13860</i> (<i>Bsal</i>)	5'-CACATGGTCTCTTCTGGTGGCAATGATC-3'
<i>EGO55_13860</i> UpRv_		5'-ACATGGTCTCGTCATATGCCCTGATTATGTC-3'
<i>EGO55_13860</i> DwFw_	Amplification of the fragment downstream <i>EGO55_13860</i> (<i>Bsal</i>)	5'-CACATGGTCTCGATGACCTTCTCTCCGAGCT-3'
<i>EGO55_13860</i> DwRv_		5'-CACATGGTCTCGTTCCCAACAACCTCTGCCTAT-3'

(Continued on next page)

TABLE 1 Bacterial strains, plasmids, and primers used in this study (Continued)

Strain, plasmid, or primer	Genotype and/or description	Reference or sequence
EGO55_13845FwXbal	Amplification of EGO55_13845 (Xbal-SphI)	5'-ATATTCTAGAAAGGAGGTGAATATGTCTCTGTGAATAACGGT-3'
EGO55_13845RvSphI		5'-ATATGCATGCC TAGCTGATGACGTAATCGG-3'
EGO55_13860FwHindIII	Amplification of EGO55_13860 (HindIII-SpeI)	5'-ATATAAGCTTAAGGAGGTGAATATGAAGTTTCCATCATTATGAA-3'
EGO55_13860RvSpeI		5'-ATATACTAGTTCAGCCAGCCACAGTTTCAA-3'
GibEGO55_13845Fw	Amplification of EGO55_13845 and cloning by Gibson assembly	5'-TAAGAAGGAGATATACATATGTCTCTGTGAATAACG-3'
GibEGO55_13845Rv		5'-GTGGTGGTGGTGGTGGTGGCTGATGACGTAATCG-3'
GibEGO55_13860Fw	Amplification of EGO55_13860 and cloning by Gibson assembly	5'-TAAGAAGGAGATATACATATGAAGTTTCCATCATTTA-3'
GibEGO55_13860Rv		5'-GTGGTGGTGGTGGTGGTGGCCAGCCACAGTTT-3'

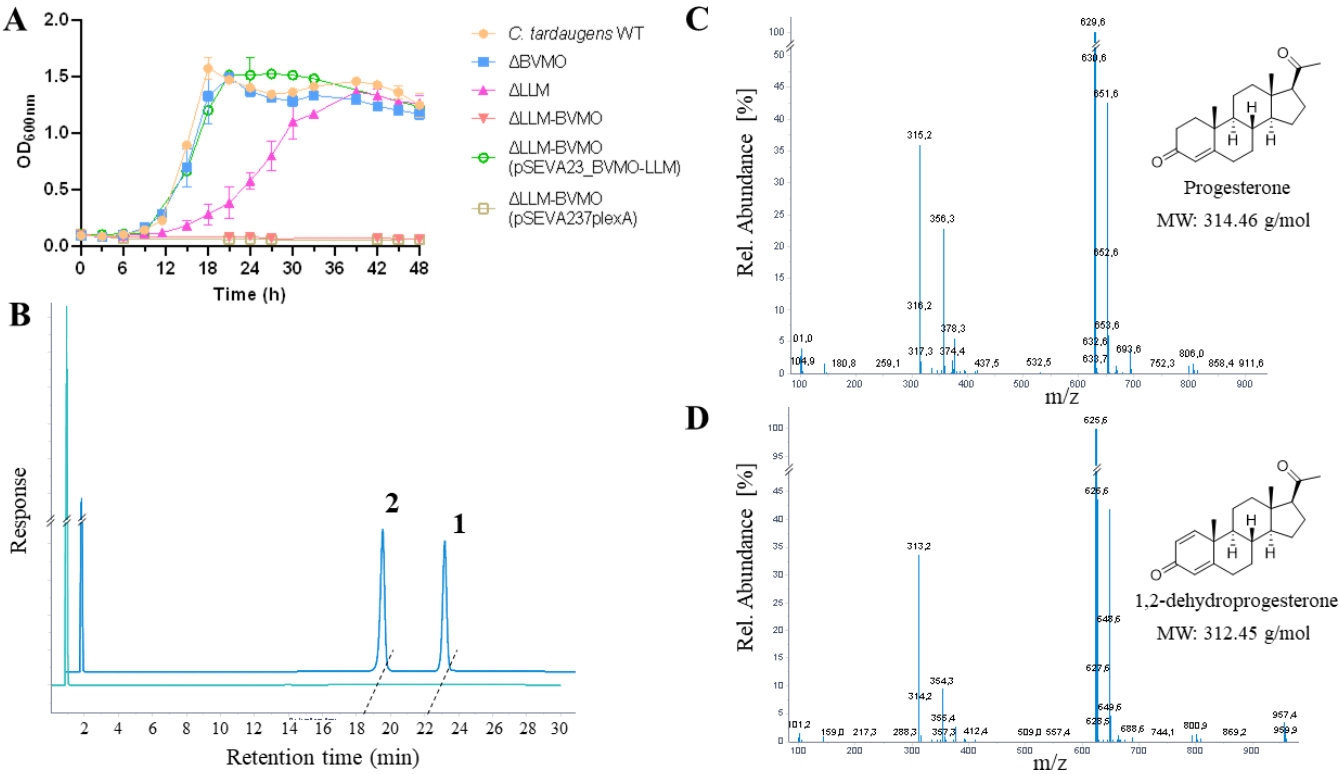


FIG 5 (A) Growth curve of *C. talarum* WT, Δ BVMO, Δ LLM, Δ LLM-BVMO, Δ LLM-BVMO (pSEVA23_BVMO-LLM), and Δ LLM-BVMO (pSEVA237plexA) strains on 1.71 mM PROG. (B) High-performance liquid chromatograph with mass spectrometry (HPLC-MS) analysis of the organic phase extracted from *C. talarum* WT (turquoise) and Δ LLM strain (blue) culture grown in progesterone for 24 h. Peak 1 corresponds to progesterone, and peak 2 corresponds to 1,2-dPROG. (C) Mass spectra of peak 1, corresponding to PROG. (D) Mass spectra of peak 2, corresponding to 1,2-dPROG.

stable throughout 96 h of cultivation (Fig. S4), demonstrating that this mutant is capable of transforming PROG into the 1,2-dPROG intermediate, but is not able to use PROG as a carbon and energy source.

BVMO and LLM catalyze the transformation of PROG into TES-Ac

To determine the precise role of BVMO and LLM enzymes in PROG catabolism, we performed *in vitro* biochemical assays with BVMO and LLM using PROG as substrate.

We initially conducted enzymatic assays using crude protein extracts from *C. talarum* WT and mutant strains cultured in PROG, but we were unable to detect the formation of TES-Ac (Fig. S5), most probably due to the low expression of the enzymes. Therefore, to further characterize their activities, we overexpressed the enzymes in a heterologous system. To this aim, the genes encoding BVMO and LLM were cloned in the pET29a(+) expression vector, yielding pET-29(+)-BVMO and pET-29(+)-LLM, respectively.

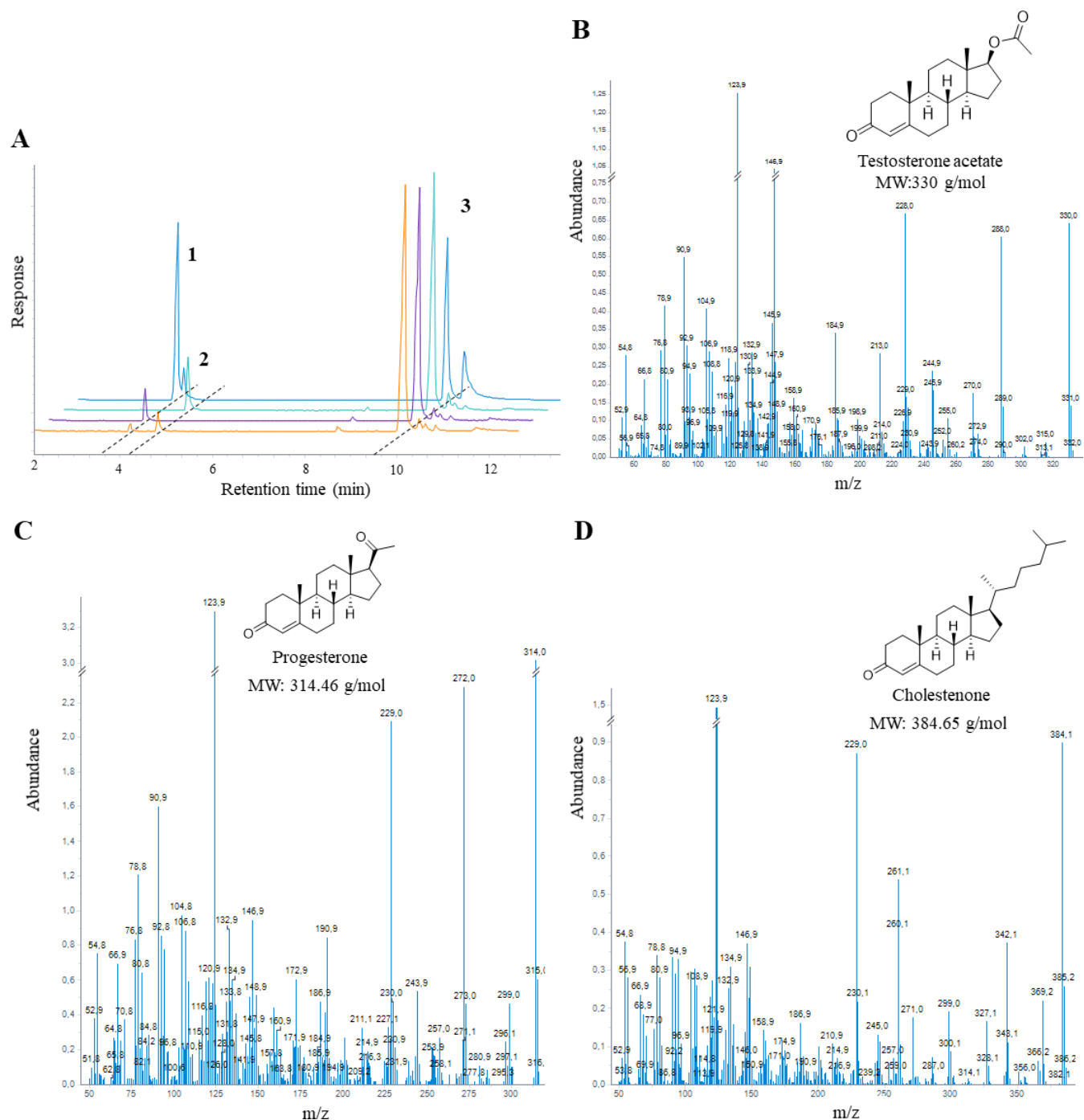


FIG 6 (A) Gas chromatography-mass spectrometry (GC-MS) analysis of the organic phase extracted from the enzymatic reaction of progesterone with crude extracts from *E. coli* BL21(DE3) (pET-29(+)/BVMO) (orange), *E. coli* BL21(DE3) (pET-29(+)/LLM) (purple) and *E. coli* BL21(DE3) [pET29(+)] (turquoise). Commercial testosterone acetate is represented in blue. Peaks correspond to (1) TES-Ac, (2) PROG, and (3) cholestenone (ISTD). (B) Mass fragmentation pattern of peak 1 in GC-MS. (C) Mass fragmentation pattern of peak 2 in GC-MS. (D) Mass fragmentation pattern of peak 3 in GC-MS.

Both plasmids were transformed in *E. coli* BL21(DE3), and the enzymes were overexpressed (Fig. S6). Crude protein extracts were used to perform an *in vitro* enzymatic assay using PROG and NAD(P)H as substrates (see Materials and Methods). After 30 min of reaction, the steroids were extracted and analyzed by gas chromatography-mass spectrometry (GC-MS) and compared to the control with the empty plasmid pET29(+). Both BVMO and LLM enzymatic assays deliver a new peak when analyzed by GC-MS (Fig.

6A). According to its fragmentation pattern and the NIST Mass Spectral Database, it was identified as TES-Ac (Fig. 6B). This result confirmed that *pd*c-encoded BVMO and LLM are able to transform PROG into TES-Ac through a Baeyer-Villiger oxidation.

BVMO and LLM exhibit differential activity on PROG derivatives

We have demonstrated that the BVMO and LLM enzymes have an important role in PROG catabolism in *C. tar*daugens. However, they exhibit redundancy in their enzymatic reactions, as both monooxygenases catalyse a Baeyer-Villiger oxidation of PROG, leading to the formation of TES-Ac. To explore new substrates for both monooxygenases and evaluate the different activities that may explain their essential role in the PROG degradation pathway, an enzymatic assay using progesterone-related steroids was conducted. Crude extracts from *E. coli* BL21(DE3) containing pET-29(+)BVMO and pET-29(+)LLM plasmids were obtained. The enzymatic activity was analyzed on four steroids: PROG as the native substrate, 1,2-dPROG as the intermediate detected in PROG degradation, 21-hydroxyprogesterone (21-OHPROG) as a steroid having a modification on the side chain, and PREG as the reduced and isomerized PROG derivative. The activity of the BVMO was assessed by measuring the decrease in absorbance of 0.2 mM NADPH at 340 nm when 0.5 mM of substrate was added. To assess the basal consumption of NADPH of the crude extract, NADPH was added with no substrate. The BVMO EGO55_13845 exhibited the highest specific activity on PROG and PREG (70.4 and 68.5 U/mg, respectively) (Table 2), indicating that, among all the substrates tested, PROG and PREG are preferred by the enzyme, with no significant difference between them. Nevertheless, when 1,2-dPROG was tested as substrate, the specific activity dropped to 41 U/mg, suggesting that the new double bond between C1 and C2 may hinder the proper positioning of this steroid in the substrate binding site. The lowest specific activity was observed in the presence of 21-OHPROG (8.7 U/mg) (Table 2), which is an expected result due to the fact that the hydroxyl group next to the ketone group may cause a steric hindrance in the protein’s active site. The specific activity of the LLM enzyme was defined as the nanomole of substrate consumed per minute per milligram of protein, as measuring NADH consumption would be an indirect method due to the involvement of a reducing partner in the reaction. The reaction was initiated by adding 5 mM NADH, and samples were collected after 5 and 30 min for analysis via HPLC-MS. LLM exhibited similar specific activity with PROG, PREG, and 1,2-dPROG (10.1, 12.5, and 14.7 U/mg, respectively) (Table 2), with 1,2-dPROG showing a slightly higher value. As observed in the BVMO assays, the specific activity of LLM on 21-OHPROG was the lowest, reaching only 6.4 U/mg. Additionally, after 30 min of incubation, PROG and PREG remained detectable in the sample, while 1,2-dPROG and 21-OHPROG were no longer detected (Fig. S7 to S9). Although these enzymatic assays with both monooxygenases were performed using crude extracts and cannot be directly compared, it is evident that each enzyme exhibits distinct substrate activities, which may explain their specific biological roles within the *pd*c cluster. While BVMO preferred substrates are PROG and PREG, LLM exhibits similar

TABLE 2 Specific activity over different substrates of crude extracts containing BVMO EGO55_13845 or LLM EGO55_13860 from *C. tar*daugens

Enzyme	Substrate	Specific activity (U/mg) ^a
BVMO EGO55_13845	Progesterone	70.4
	21-Hydroxyprogesterone	8.7
	1,2-Dehydroprogesterone	41
	Pregnenolone	68.5
LLM EGO55_13860	Progesterone	10.1
	21-Hydroxyprogesterone	6.4
	1,2-Dehydroprogesterone	14.7
	Pregnenolone	12.5

^aFor BVMO, specific activity is defined as the amount of NADPH (nmol) consumed per minute per milligram of protein. For LLM, specific activity is defined as the number of nmol of substrate consumed per minute per milligram of protein.

activity on PROG, PREG, and 1,2-dPROG, suggesting that the substrate binding domain of LLM can accommodate steroids with structural modifications in ring A. Both monooxygenases share the common trait that any modification in the side chain of PROG (e.g., hydroxylation) hinders their activity. In order to precisely identify the product obtained in the enzymatic assays, the samples were analyzed by GC-MS. When 1,2-dPROG was used as substrate, both monooxygenases produced boldenone acetate (BOL-Ac) (Fig. S7). PREG was transformed into androstenediol 17-acetate (Fig. S8), and despite a low specific activity, 21-OHPROG was oxidized into testosterone 2-hydroxyacetate (Fig. S9).

These data suggest that BVMO and LLM are induced in order to catalyze a Baeyer-Villiger oxidation on PROG to produce TES-Ac *in vivo*. However, when PROG enters the cell, a fraction of PROG can be dehydrogenated, before being transformed into TES-Ac, by a 3-ketosteroid- Δ^1 -dehydrogenase (KstD) to yield 1,2-dPROG, requiring the involvement of enzymes capable of performing monooxygenation on both substrates, PROG and 1,2-dPROG. We suggest that in *C. tardaugs*, 1,2-dPROG monooxygenation activity might be supplied mainly by LLM EGO55_13860. This hypothesis is supported by the longer lag phase observed in the Δ LLM mutant strain growing on PROG (Fig. 5A), where the accumulation of 1,2-dPROG is observed up to 39 h of growth (Fig. S2).

C. tardaugs PROG metabolism is linked to TES metabolism

Taking all these results into account, PROG catabolism in *C. tardaugs* likely begins with the transformation of PROG into TES-Ac, enabling its entry into the 9,10-seco pathway (20) (Fig. 7). However, we have to consider that PROG can also be dehydrogenated by a KstD to form 1,2-dPROG, an enzyme widely distributed in steroid-degrading bacteria (49–51). KstDs are deeply studied, and mutants in their encoding genes are currently used to accumulate C19 and C21 steroids as pharmaceutical synthons (52–55).

Then, PROG and 1,2-dPROG can undergo a Baeyer-Villiger monooxygenation performed by either BVMO EGO55_13845 or LLM EGO55_13860 to form TES-Ac or BOL-Ac, respectively. The fact that 1,2-dPROG is an intermediate in PROG metabolism

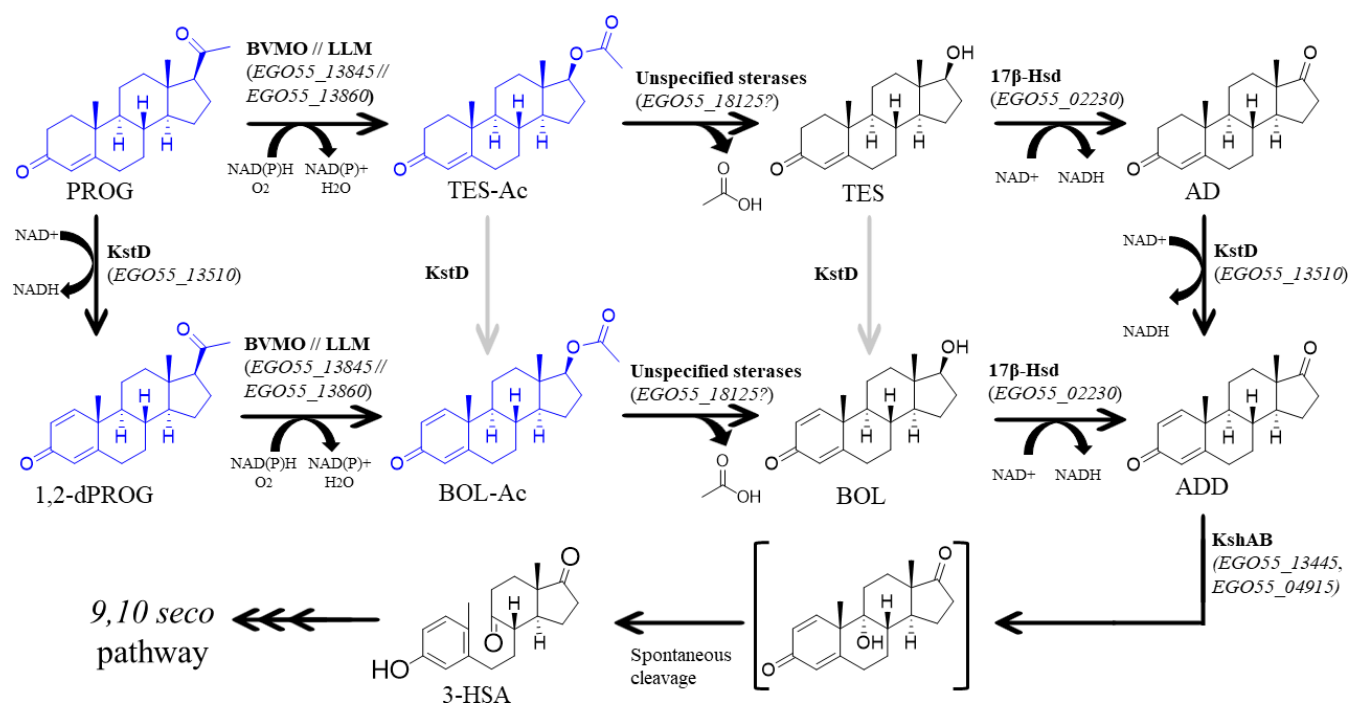


FIG 7 Proposed progesterone degradation pathway in *C. tardaugs* NBRC 16725. Enzymes catalyzing the reaction are marked in bold, and their proposed encoding genes are specified in brackets. The compounds detected in this work are colored in blue. Compound names are indicated with an abbreviation: 3-HSA, 3-hydroxy-9,10-secoandrost-1,3,5(10)-triene-9,17-dione. Enzyme names are BVMO; LLM; 17β-Hsd, 17β-hydroxysteroid dehydrogenase; KshAB, 3-ketosteroid 9α-hydroxylase.

does not necessarily imply that the first step of PROG degradation is a C1 dehydrogenation. The C1 dehydrogenation by KtsD enzymes could also occur after the Bayer-Villiger reaction on PROG. In fact, since the accumulated 1,2-dPROG disappears over time, this suggests that the Bayer-Villiger reaction can also occur on 1,2-dPROG, leading to the formation of BOL-Ac instead of TES-Ac.

The transformation of TES-Ac or BOL-Ac into TES or boldenone (BOL) could be carried out by an unknown sterol acetyl hydrolase (56). In this sense, the gene *EGO55_18125* from *C. talarum* is annotated as an acetyl hydrolase. The transcriptomic analysis indicates that this gene is neither induced nor repressed, since its TPMs are between 7 and 9 in both PROG and TES conditions. Nevertheless, lipases have also been described to hydrolyze TES esters (57), and we cannot rule out the possibility that lipases or other hydrolases in *C. talarum* may play this role.

Once TES or BOL are formed, they enter the TES degradation pathway previously described by Ibero et al. in *C. talarum*, where TES and BOL are oxidized into AD and ADD, respectively, by the 3,17 β -hydroxysteroid dehydrogenase encoded by the *EGO55_02230* gene (20). AD is dehydrogenated into ADD by a KstD, which is further hydroxylated by a 3-ketosteroid 9 α -hydroxylase that opens the steroidal ring A, facilitating its subsequent mineralization via the 9,10-seco pathway.

DISCUSSION

In spite of the fact that PROG is a critical EDC contaminant, the genes involved in its degradation pathway in bacteria have not been elucidated yet, and the few studies concerning this topic only describe some of the possible intermediates of its degradation pathway. Thus, when Liu et al. studied the aerobic degradation of PROG and norgestrel by bacteria from activated sludge, they detected AD, ADD, 17 β -boldenone, TES, 6,7-dehydroprogesterone, and 1,2-dPROG as intermediates in the culture medium (28). They proposed that PROG undergoes different transformations ending up in ADD or 17 β -boldenone, which are finally mineralized. Horinouchi et al. also suggested that PROG was transformed somehow into AD and then degraded through the well-described 9,10-seco pathway, but the genes that channel PROG into this degradative pathway were not described (20). Moreover, although a BVMO able to transform PROG into TES-Ac has been isolated in *R. rhodochrous* (34), the PROG catabolic pathway was not described in this microorganism.

Taking advantage of the fact that *C. talarum* is able to efficiently use many different steroids including PROG and that we had developed the genetic tools to manipulate this microorganism, we have investigated the fundamental steps involved in the degradation of PROG in bacteria. Because we assumed that most probably the catabolism of PROG and TES was related, we performed a differential transcriptomic analysis of the cells cultured in PROG or TES. This analysis revealed the upregulation of the *pdc* cluster encoding two monooxygenases: a type I BVMO encoded by *EGO55_13845* and an LLM encoded by *EGO55_18860* that transform PROG into TES-Ac through a Bayer-Villiger reaction.

As previously mentioned, BVMOs are a well-known family of proteins that perform a characteristic monooxygenation on a wide range of substrates with a ketone group such as phenylacetone, 2-octanone, cyclobutanone, and different steroids (58, 59). Kólek et al. studied the D-lactonization catalyzed by a BVMO from *Penicillium camemberti* AM83, which transforms steroids such as PROG, TES, and DHEA into testololactone (59). They proposed this fungus as a potential biocatalyst for producing steroidal lactones with novel biologically active applications. The presence of BVMO was postulated as well by Javid et al., showing the bioconversion of progesterone to testololactone by the fungus *Aspergillus sojae* PTCC 5196 through a three-step pathway (17 β -acetyl side chain cleavage, 17 β -hydroxyl oxidation, and oxygenative lactonization of 17-ketone) (30). The specificity of a BVMO from the bacterium gram-positive *Dietzia* sp. D5 has also been tested and resulted in a wide substrate range including cyclobutanone, phenylacetone, PROG, and even sulfides such as thionisole and ethionamide (58). Moreover, the 3D

structure of a BVMO from *R. rhodochrous* transforming PROG into TES-Ac has been solved, leading to the identification of the FAD- and NADP-binding domains and the substrate binding site (33).

On the other hand, LLM enzymes are less studied, likely due to the complexity of their catalysis, which involves a complex of two proteins: a monooxygenase and a reductase. The only known examples are the AbmE2/AbmZ complex from *Streptomyces koyagensis* SCSIO 5802, which catalyzes the inactivation of the antimicrobial abyssomicin 2 to neoabyssomicin B (41), and the 2,5-diketocamphane 1,2-monooxygenases and 3,6-diketocamphane 1,6-monooxygenases, along with their respective reductases, involved in the camphor degradation pathway of *Pseudomonas putida* ATCC 17453 (60, 61).

The functional analyses performed through mutagenesis and subsequent complementation demonstrated that both enzymes are functionally redundant, meaning they can catalyze the same reaction. As a result, both enzymes had to be deleted in order to create a mutant incapable of growing on PROG as its sole carbon and energy source. The functional analyses carried out with the mutant strains showed that PROG can be transformed into 1,2-dPROG most probably due to the dehydrogenation activity of KstD, an enzyme widely distributed in steroid-degrading bacteria (49–51). KstDs are deeply studied, and mutants in their encoding genes are currently used to accumulate C19 and C21 steroids as pharmaceutical synthons (52–55, 62). In a previous work, we demonstrated that the *EGO55_13510* gene encodes a KstD involved in TES degradation in *C. tardaugens* (20). However, other putative *kstD* genes, such as *EGO55_01175* and *EGO55_03150*, are also present within its genome and might play a role in the transformation of PROG into 1,2-dPROG. In this sense, KstDs are known to act on steroid substrates with or without different side chains in C17, since C1 dehydrogenation is a fundamental step in steroid catabolism to facilitate aromatization of the A ring (63). The transcriptomic analysis showed that all *kstD* genes had an FC between –1 and 1 and a TPM between 2 and 150 (Table S1). This suggests that these genes are likely equally expressed under both TES and PROG conditions. The finding that 1,2-dPROG is an intermediate in PROG metabolism does not necessarily indicate that the first step of PROG degradation is C1 dehydrogenation. This is because C1 dehydrogenation by KstD enzymes could also occur after the Baeyer-Villiger reaction on PROG. In any case, since 1,2-dPROG disappears over time, it suggests that the Baeyer-Villiger reaction can also take place on 1,2-dPROG, leading to the formation of BOL-Ac instead of TES-Ac. The presence of functionally redundant BVMO and LLM enzymes in *C. tardaugens* can be explained by their distinct substrate preferences. Each enzyme may act differentially on steroids other than PROG, such as 1,2-dPROG, contributing to the versatility of the organism's steroid degradation pathway.

The search for the *pdc* cluster in other bacteria opens up the possibility of identifying microorganisms whose ability to degrade PROG has not yet been described. This is the case of *C. estronivorus*, formerly known as *Altererythrobacter estronivorus*, which is a gram-negative bacterium known to degrade estrogens, although its growth in PROG has not been tested (64). *N. aromaticivorans* is a widely studied strain due to its versatile metabolism, being able to degrade lignin aromatic derivatives and produce valuable compounds. However, its growth in different steroids has not been assessed yet (65–67). The case of *C. testosteroni* is different because it has been demonstrated that it can grow on TES, PROG, and other steroids like cholic acid as carbon and energy source, but the genes involved in PROG degradation had not been identified (68). Remarkably, the *pdc* homologous genes of *N. aromaticivorans* are located in the 487 kb plasmid pNL2, but the BVMO (*Saro_3708*) and TetR-like regulator (*Saro_3709*) are 123 kb separated from the LLM (*Saro_3595*) and the LuxR-like regulator (*Saro_3594*) (Fig. 4). Therefore, this finding suggests that both putative Baeyer-Villiger enzymes have not evolved together in this bacterium and might play different functions. Moreover, both regulators might be involved in controlling the expression of the respective accompanying genes. Why these genes are located contiguously in the genomes of *C. tardaugens* and *C.*

estronivorus cannot be easily explained and requires further studies. This comparative analysis provides an interesting observation concerning the role of flavin reductases located close to the *pdC* cluster, since a flavin reductase was not found near the LLM of *N. aromaticivorans* nor the LLM of *N. arvorzyae* (Fig. 4), suggesting that these LLM enzymes could be assisted by other flavin reductases located anywhere else in the genome.

MATERIALS AND METHODS

Chemicals

PROG, 21-OHPROG, PREG, TES, TES-Ac, TES-Cyp, AD, ADD, DHEA, E2, sodium cholate, sodium deoxycholate, chloroform, *n*-hexane, ethyl acetate, and acetonitrile were purchased from Merck KGaA Sigma (Germany). Randomly methylated β -cyclodextrin (CDX) was purchased from Cyclolab R&D Ltd. (Hungary). 1,2-dPROG was purchased from Toronto Research Chemicals (Canada).

Strains and growth media

All bacterial strains, plasmids, and primers used in this study are listed in Table 1. *C. tarchaogens* NBRC 16725 (formerly known as *N. tarchaogens*) was obtained from the Leibniz-Institut DSMZ-type culture collection. This strain and all its mutants were cultured as described (20). *C. tarchaogens* was cultured in NB (Difco) as rich medium at 30°C and 200 rpm in an orbital shaker, while the minimal medium used was M63 (KH₂PO₄ [136 g/L], (NH₄)₂SO₄ [20 g/L], FeSO₄·7H₂O [5 mg/L], pH 7.0) supplemented with 0.39 mM CaCl₂, 1 mM MgSO₄, and the appropriate carbon source concentration. Steroids stock solutions were prepared in phosphate-buffered saline (PBS) buffer (per liter, 8 g NaCl, 0.2 g KCl, 1.44 g Na₂HPO₄, and 0.24 g KH₂PO₄; pH 6.8) with 70 mM CDX, and the final concentration in the culture was 13.33 mM CDX and 1.71 mM and 1.89 mM for PROG and TES, respectively. *Escherichia coli* DH10B, *E. coli* BL21(DE3), and *E. coli* HB101 strains were grown in lysogeny broth (LB) medium at 37°C and 200 rpm in an orbital shaker. When needed, rifampicin (50 µg/mL) and kanamycin (50 µg/mL or 10 µg/mL) were added for *E. coli* and *C. tarchaogens* strains, respectively.

RNA extraction

Total RNA extraction of *C. tarchaogens* was performed as described (20). Cells from three biological replicates were first cultured on minimal medium with PROG or TES as carbon sources until they reached the mid-exponential phase (OD₆₀₀ = 0.6), when they were harvested and stored at -80°C. Pellets were then thawed and lysed in 400 µL TE buffer (10 mM Tris-HCl, 1 mM EDTA, pH 7.5) with lysozyme (50 mg/mL). Three freezing-thawing cycles were performed before using the High Pure Isolation Kit (Roche, Switzerland), followed by DNA-free DNA Removal Kit (Invitrogen, UK) treatment to obtain pure RNA. Purity and concentration were measured in a ND1000 spectrophotometer (Nanodrop Technologies, USA), while RNA integrity was checked in an Agilent Technologies 2100 Bioanalyzer (USA).

Transcriptomic analysis (RNA-seq)

Transcriptome sequencing was done by MacroGen NGS Service, using an Illumina TruSeq RNA library with 6GB/sample sequencing coverage, obtaining fragments of 151 bp paired-end reads. Bioinformatics analyses were performed by the Bioinformatics and Biostatistics Service of the Center for Biological Research Margarita Salas (CIBMS-CSIC). Raw reads data quality was checked using FastQC, and then they were trimmed and cleaned with Trimmomatic 0.39 (69). After filtration, 55,475,633 trimmed reads were obtained, and 45,845,222 million high-quality clean reads were mapped to the genome of *C. tarchaogens* (accession number CP034179) using Bowtie2 2.4.2 (70). Expression quantification was done using HTSeq-count 0.13.5 (71), and differential gene expression analysis was performed using DESeq2 1.32.0 from the R software v.4.3.2. (R: The R Project

for Statistical Computing; <https://www.r-project.org/>). Genes presenting an $FC > 2$ or $FC < -2$ and false discovery rate < 0.05 (represented as *padj* in Table S1) were considered as differentially expressed. In the dissimilarity matrix shown in the heatmap, the color scale represents the $\log_2$ of median of ratios normalized counts plus one, and the dendrogram was obtained with the Euclidean distance between genes and the complete linkage hierarchical method for clustering. The figures were done with the package ComplexHeatmap (72) in the R software v.4.3.2, indicated before.

Construction of *C. tarдаugens* knockout strains and their complementation

Knockout strains of *C. tarдаugens* were constructed by double homologous recombination as described (20). Briefly, an upstream and a downstream region of about 700 bp flanking the gene to delete were amplified and cloned using *BsaI* in the suicide vector pK18msg (Table 1), a variant of pK18mobsacB where the multiple cloning site (MCS) was substituted by type II restriction enzymes kindly provided by Godoy et al. (46). The recombinant vector was transformed into *E. coli* DH10B competent cells as described (73) and checked by PCR and by sequencing of the cloned region. Then, *C. tarдаugens* rifampicin resistant was transformed with the suicidal plasmid by triparental conjugation using *E. coli* HB101 (pRK600) as helper and *E. coli* DH10B, harboring the corresponding vector, as donor (47, 74). The result of the conjugation was plated on NB agar with 10 $\mu\text{g/mL}$ kanamycin and 50 $\mu\text{g/mL}$ rifampicin. Candidates were checked by colony PCR to ensure that the suicide vector is inserted in the genome. Then, they were cultured to stationary phase and plated on NB agar with 5% sucrose. Only the clones that were able to grow in the presence of sucrose and sensitive to kanamycin were checked by PCR in order to confirm the complete deletion of the gene after the second recombination event.

C. tarдаugens mutants were complemented with plasmids belonging to the pSEVA collection (pSEVA23) by triparental conjugation as described above. The genes *EGO_13845* and *EGO55_13860*, previously deleted, were amplified using the primers indicated in Table 1, cloned in *XbaI*-*SphI* and *HindIII*-*SpeI* sites, respectively, and expressed under the *P_{lexA}* constitutive promoter expression (48).

Cloning and expression of *EGO55_13845* and *EGO55_13860*

EGO55_13845 and *EGO55_13860* genes, encoding a BVMO and LLM, respectively, were amplified with primers described in Table 1 and cloned separately in the pET-29a(+) vector using the Gibson Assembly Cloning Kit (New England Biolabs) following the manufacturer instructions. The resulting pET-BVMO and pET-LLM plasmids were then electroporated in *E. coli* BL21(DE3) electrocompetent cells using a Gene Pulser (Bio-Rad, USA) (75). Transformed strains were cultured in 50 mL of LB with kanamycin (50 $\mu\text{g/mL}$) up to an OD_{600} of 0.5–0.8 at 37°C and 200 rpm. Then, isopropyl β -D-1-thiogalactopyranoside (IPTG) was added to a final concentration of 0.2 mM to induce gene expression. Cultures were then kept at 30°C for 3 h. Next, cells were harvested and washed twice with 0.85% saline solution and finally resuspended in 50 mM phosphate buffer (pH 8.0). A Branson sonicator (Germany) was used to lyse the cells with three cycles of 10 s at maximum power and 30 s cooling on ice. Crude extract was centrifuged at $13,000 \times g$ in a microfuge Sigma 1-15K (Germany) for 15 min at 4°C to obtain the soluble fraction. Protein concentration was calculated with the Bradford method (76), and overproduction of proteins was visualized by SDS-polyacrylamide gel electrophoresis (Fig. S6).

Enzymatic assay of BVMO and LLM activities

BVMO activity was assayed in a Shimadzu UV-1900i spectrophotometer (Japan) by measuring the decrease in absorbance of NADPH at 340 nm ($\epsilon_{\text{NADPH}} = 6.22 \text{ cm}^{-1} \text{ mM}^{-1}$) when the substrate was added. The enzymatic reaction was done in quartz cuvettes (HELLMA Analytics, 10 mm) containing 0.6 mL of 2 mg/mL of crude extract in 50 mM sodium phosphate buffer (pH 8.0) at 25°C. Substrate stocks were prepared in methanol at

50 mM to obtain a final concentration of 0.5 mM, and the reaction started when 0.2 mM NADPH was added. In this case, 1 unit of specific activity (U/mg) was defined as the number of nanomoles of NADPH consumed in 1 min per milligram of protein.

The LLM enzymatic assays were performed as follows: 0.6 mL of 4 mg/mL of crude extract in 50 mM sodium phosphate buffer (pH 7.5) was added to a 4 mL Clear Vial (Supelco, USA) with magnetic agitation at 25°C. Substrate stocks were prepared as described before and added at a final concentration of 0.5 mM, along with 20 µM of FMN. Reaction started when 5 mM NADH was added. A total of 200 µL samples was taken at 0, 5, and 30 min and mixed with 500 µL of chloroform to stop the reaction. Steroids were extracted and analyzed as described below. Since product standards were not available, 1 unit of specific activity (U/mg) was defined as the number of nanomoles of substrate consumed in 1 min per milligram of protein.

To determine the enzymatic activities in protein crude extracts of *C. tardaugens*, the WT and mutant strains were grown in 12.5 mL of NB in the presence of 1.71 mM PROG to ensure *EGO55_13845* and *EGO55_13860* induction. When the culture reached the mid-exponential phase ($OD_{600} = 1$), cells were harvested and washed twice with 0.85% saline solution. Finally, cells were resuspended in 50 mM phosphate buffer (pH 8.0) and sonicated as described before. Crude extract was centrifuged at $13,000 \times g$ in a microfuge Sigma 1-15K (Germany) for 15 min at 4°C to obtain the soluble fraction. Thereafter, 0.6 mL of protein crude extracts was tested in a 4 mL Clear Vial (Supelco, USA) at a final concentration of 4 mg/mL. PROG was added at a final concentration of 0.5 mM with 20 µM FMN in 50 mM sodium phosphate buffer (pH 7.5). The reaction conditions were maintained at 25°C with magnetic agitation. Reaction started when 2 mM of NADPH or NADH was added. Samples of 200 µL were taken at 0 and 30 min of reaction time for their analyses.

Organic phase extraction and chromatographic analysis

The steroidal compounds from the culture media and from the enzymatic assays were extracted by organic solvent extraction with chloroform as described before (20). Briefly, two volumes of chloroform were added to the sample, vortexed, and then the organic phase was extracted, dried, and resuspended in 150 µL of acetonitrile. Extracted steroids were detected and identified using high-efficiency liquid chromatography coupled to a photodiode array detector and mass spectroscopy (HPLC-MS) as described in Hernández-Fernández et al. (54). Samples were analyzed in a 1260 Infinity II (Agilent, USA) with InfinityLab LC/MSD XT (Agilent, USA), with automatic injector, diode array detector, and ion trap with electrospray. For the separation method, a Poroshell 120 EC-C18 (4 µm; 4.6×100 mm) (Agilent, USA) column was used, with a flow rate of 1 mL/min starting with 70% of solvent A ($H_2O + 0.1\%$ formic acid) and 30% of solvent B (acetonitrile + 0.1% formic acid) for 5 min, followed by an increase of solvent B to 55% for 20 min and then reduced back to 30% for 2 min. This rate was finally maintained for 3 min. The elution was monitored at 195, 210, 230, 250, and 275 nm, and the steroid compounds obtained were identified, thanks to the commercial standards comparing mass spectra and retention time.

To identify the steroid compounds, a GC-MS analysis was performed as described (55). Previously to the steroid extraction, 25 µL of cholestenone 10 mM was added to each sample as an internal standard. Once extracted, trimethylsilyl ether derivatives were formed by adding 50 µL of N,O-bis(trimethylsilyl)trifluoroacetamide with trimethylchlorosilane (Merck, Germany) and 50 µL of pyridine. This solution was then heated at 80°C for 20 min. An Agilent 7890A gas chromatograph coupled to an Agilent 5975 C mass detector (Agilent Technologies, USA) was used for the GC-MS analysis. Mass spectra were recorded in electron impact mode at 70 eV within an m/z range set between 50 and 550. A 30 m $\times$ 0.25 mm i.d. capillary column (0.25 µm film thickness) HP-5MS (5% diphenyl, 95% dimethylpolysiloxane) (Agilent Technologies, USA) was used with the chromatograph, and working conditions were a split ratio of 20:1, 320°C for the injector temperature. Then the column temperature was set at 240°C for 3 min and finally heated

to 320°C at 5°C min⁻¹. In order to identify the compounds obtained, mass spectra and retention data were compared to those observed in the commercial standards and to those present in the NIST Mass Spectral Database (NIST 2011).

ACKNOWLEDGMENTS

The authors declare that financial support was received for the research, authorship, and/or publication of this article. This work was supported by grants from the Spanish Ministry of Science and Innovation PID2021-125370OB-I00. The grant for one of the authors (G.H.-F.) with reference FPU21/05101 is gratefully recognized.

B.G.: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Project administration, Resources, Supervision, Writing – original draft, Writing – review & editing. J.L.G.: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Project administration, Resources, Supervision, Writing – original draft, Writing – review & editing. J.I.: Data curation, Formal analysis, Investigation, Methodology. G.H.-F.: Data curation, Formal analysis, Investigation, Methodology, Writing – original draft, Writing – review & editing.

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations.

AUTHOR AFFILIATIONS

¹Department of Biotechnology, Centro de Investigaciones Biológicas Margarita Salas (CSIC), Madrid, Community of Madrid, Spain

²Department of Biochemistry and Molecular Biology, Chemical Sciences Faculty, Complutense University of Madrid, Madrid, Community of Madrid, Spain

AUTHOR ORCIDs

Gabriel Hernández-Fernández  <http://orcid.org/0000-0002-7420-2313>

Beatriz Galán  <http://orcid.org/0000-0002-2596-6034>

FUNDING

Funder	Grant(s)	Author(s)
Ministerio de Economía y Competitividad	PID2021 - 125370OB - I00	Gabriel Hernández-Fernández Juan Ibero José L. García Beatriz Galán
Ministerio de Economía y Competitividad	FPU21/05101	Gabriel Hernández-Fernández

AUTHOR CONTRIBUTIONS

Gabriel Hernández-Fernández, Data curation, Formal analysis, Investigation, Methodology, Validation, Writing – original draft, Writing – review and editing | Juan Ibero, Data curation, Formal analysis, Investigation, Methodology, Validation | José L. García, Conceptualization, Funding acquisition, Resources, Supervision, Writing – original draft, Writing – review and editing | Beatriz Galán, Conceptualization, Funding acquisition, Investigation, Resources, Supervision, Writing – original draft, Writing – review and editing

DATA AVAILABILITY

Raw read data obtained from the three replicates of the transcriptome of the strain grown on PROG and TES have been deposited in the Sequence Read Archive (SRA) database of the National Center for Biotechnology Information (NCBI) under accession numbers [SRR31340002](#) and [SRR31340003](#) (Bioproject [PRJNA1185793](#)), respectively. The original contributions presented in this study are included in this article and supplemental material; further inquiries can be directed to the corresponding author.

ADDITIONAL FILES

The following material is available [online](#).

Supplemental Material

Supplemental figures (mBio01077-25-s0001.pdf). Figures S1 to S9.

Table S1 (mBio01077-25-s0002.xlsx). Transcriptomic analysis of *C. tauradgens* in progesterone versus testosterone.

REFERENCES

- Barbosa MO, Moreira NFF, Ribeiro AR, Pereira MFR, Silva AMT. 2016. Occurrence and removal of organic micropollutants: an overview of the watch list of EU Decision 2015/495. *Water Res* 94:257–279. <https://doi.org/10.1016/j.watres.2016.02.047>
- Milla S, Depiereux S, Kestemont P. 2011. The effects of estrogenic and androgenic endocrine disruptors on the immune system of fish: a review. *Ecotoxicology* 20:305–319. <https://doi.org/10.1007/s10646-010-0588-7>
- Teles M, Gravato C, Pacheco M, Santos MA. 2004. Juvenile sea bass biotransformation, genotoxic and endocrine responses to beta-naphthoflavone, 4-nonylphenol and 17 beta-estradiol individual and combined exposures. *Chemosphere* 57:147–158. <https://doi.org/10.1016/j.chemosphere.2004.02.023>
- Fan Z, Wu S, Chang H, Hu J. 2011. Behaviors of glucocorticoids, androgens and progestogens in a municipal sewage treatment plant: comparison to estrogens. *Environ Sci Technol* 45:2725–2733. <https://doi.org/10.1021/es103429c>
- Fan Z, Casey FXM, Hakk H, Larsen GL. 2007. Persistence and fate of 17 β -estradiol and testosterone in agricultural soils. *Chemosphere* 67:886–895. <https://doi.org/10.1016/j.chemosphere.2006.11.040>
- Kolatorova L, Vitku J, Suchopar J, Hill M, Parizek A. 2022. Progesterone: a steroid with wide range of effects in physiology as well as human medicine. *Int J Mol Sci* 23:7989. <https://doi.org/10.3390/ijms23147989>
- Golovko O, Šauer P, Fedorova G, Kroupová HK, Grabic R. 2018. Determination of progestogens in surface and waste water using SPE extraction and LC-APCI/APPI-HRPS. *Sci Total Environ* 621:1066–1073. <https://doi.org/10.1016/j.scitotenv.2017.10.120>
- Kasambala HR, Rwiza MJ, Mdegela RH. 2019. Levels and distribution of progesterone in receiving waters and wastewaters of a growing urban area. *Water Sci Technol* 80:1107–1117. <https://doi.org/10.2166/wst.2019.350>
- Sangster JL, Ali JM, Snow DD, Kolok AS, Bartelt-Hunt SL. 2016. Bioavailability and fate of sediment-associated progesterone in aquatic systems. *Environ Sci Technol* 50:4027–4036. <https://doi.org/10.1021/acs.est.5b06082>
- Cardoso PG, Rodrigues D, Madureira TV, Oliveira N, Rocha MJ, Rocha E. 2017. Warming modulates the effects of the endocrine disruptor progestin levonorgestrel on the zebrafish fitness, ovary maturation kinetics and reproduction success. *Environ Pollut* 229:300–311. <https://doi.org/10.1016/j.envpol.2017.05.090>
- Ojogboro JO, Scrimshaw MD, Sumpter JP. 2021. Steroid hormones in the aquatic environment. *Sci Total Environ* 792:148306. <https://doi.org/10.1016/j.scitotenv.2021.148306>
- Zeilinger J, Steger-Hartmann T, Maser E, Goller S, Vonk R, Länge R. 2009. Effects of synthetic gestagens on fish reproduction. *Environ Toxicol Chem* 28:2663–2670. <https://doi.org/10.1897/08-485.1>
- McCurry MD, D'Agostino GD, Walsh JT, Bisanz JE, Zalosnik I, Dong X, Morris DJ, Korzenik JR, Edlow AG, Balskus EP, Turnbaugh PJ, Huh JR, Devlin AS. 2024. Gut bacteria convert glucocorticoids into progestins in the presence of hydrogen gas. *Cell* 187:2952–2968. <https://doi.org/10.1016/j.cell.2024.05.005>
- Coombes Z, Yadav V, McCoubrey LE, Freire C, Basit AW, Conlan RS, Gonzalez D. 2020. Progestogens are metabolized by the gut microbiota: implications for colonic drug delivery. *Pharmaceutics* 12:760. <https://doi.org/10.3390/pharmaceutics12080760>
- Chiang Y-R, Wei ST-S, Wang P-H, Wu P-H, Yu C-P. 2020. Microbial degradation of steroid sex hormones: implications for environmental and ecological studies. *Microb Biotechnol* 13:926–949. <https://doi.org/10.1111/1751-7915.13504>
- Pratish A, Ye X, Yang Q, Kan J, Peng T, Wang H, Huang T, Xiong G, Hu Z. 2020. Biotransformation strategies for steroid estrogen and androgen pollution. *Appl Microbiol Biotechnol* 104:2385–2409. <https://doi.org/10.1007/s00253-020-10374-9>
- Bergstrand LH, Cardenas E, Holert J, Van Hamme JD, Mohn WW. 2016. Delineation of steroid-degrading microorganisms through comparative genomic analysis. *mBio* 7:e00166-16. <https://doi.org/10.1128/mBio.00166-16>
- Holert J, Cardenas E, Bergstrand LH, Zaikova E, Hahn AS, Hallam SJ, Mohn WW. 2018. Metagenomes reveal global distribution of bacterial steroid catabolism in natural, engineered, and host environments. *mBio* 9:e02345-17. <https://doi.org/10.1128/mBio.02345-17>
- Olivera ER, Luengo JM. 2019. Steroids as environmental compounds recalcitrant to degradation: genetic mechanisms of bacterial biodegradation pathways. *Genes (Basel)* 10:512. <https://doi.org/10.3390/genes10070512>
- Horinouchi M, Hayashi T, Kudo T. 2012. Steroid degradation in *Comamonas testosteroni*. *J Steroid Biochem Mol Biol* 129:4–14. <https://doi.org/10.1016/j.jsbmb.2010.10.008>
- Ibero J, Galán B, Díaz E, García JL. 2019. Testosterone degradative pathway of *Novosphingobium tardagens*. *Genes (Basel)* 10:871. <https://doi.org/10.3390/genes10110871>
- Ibero J, Galán B, Rivero-Buceta V, García JL. 2020. Unraveling the 17 β -estradiol degradation pathway in *Novosphingobium tardagens* NBRC 16725. *Front Microbiol* 11:588300. <https://doi.org/10.3389/fmicb.2020.588300>
- Miao L, Sun S, Ma T, Abdelrahman Yousif Abdellah Y, Wang Y, Mi Y, Yan H, Sun G, Hou N, Zhao X, Li C, Zang H. 2023. A novel estrone degradation gene cluster and catabolic mechanism in *Microbacterium oxydans* ML-6. *Appl Environ Microbiol* 89:e01489-22. <https://doi.org/10.1128/aem.01489-22>
- Zhang S, Ye X, Lin X, Zeng X, Meng S, Luo W, Yu F, Peng T, Huang T, Li J, Hu Z. 2024. Novel insights into aerobic 17 β -estradiol degradation by enriched microbial communities from mangrove sediments. *J Hazard Mater* 465:133045. <https://doi.org/10.1016/j.jhazmat.2023.133045>
- Donova MV. 2007. Transformation of steroids by actinobacteria: a review. *Prikl Biokhim Mikrobiol* 43:5–18. <https://doi.org/10.1134/S0003683807010012>

26. Donova MV, Egorova OV. 2012. Microbial steroid transformations: current state and prospects. *Appl Microbiol Biotechnol* 94:1423–1447. <https://doi.org/10.1007/s00253-012-4078-0>
27. Fernández de Las Heras L, García Fernández E, María Navarro Llorens J, Perera J, Drzyzga O. 2009. Morphological, physiological, and molecular characterization of a newly isolated steroid-degrading actinomycete, identified as *Rhodococcus ruber* strain Chol-4. *Curr Microbiol* 59:548–553. <https://doi.org/10.1007/s00284-009-9474-z>
28. Liu S, Ying G-G, Liu Y-S, Peng F-Q, He L-Y. 2013. Degradation of norgestrel by bacteria from activated sludge: comparison to progesterone. *Environ Sci Technol* 47:10266–10276. <https://doi.org/10.1021/es304688g>
29. Abdel-Kareem MM, Zohri A-N, Rasmey A-H, Hawary H. 2024. Enhancing the biotransformation of progesterone to the anticancer compound testolactone by *Penicillium chrysogenum* Ras3009: kinetic modelling and efficiency maximization. *BMC Biotechnol* 24:73. <https://doi.org/10.1186/s12896-024-00896-9>
30. Javid M, Nickavar B, Vahidi H, Faramarzi MA. 2018. Baeyer-Villiger oxidation of progesterone by *Aspergillus sojae* PTCC 5196. *Steroids* 140:52–57. <https://doi.org/10.1016/j.steroids.2018.07.008>
31. Miyamoto M, Matsumoto J, Iwata T, Itagaki E. 1995. Bacterial steroid monooxygenase catalyzing the Baeyer-Villiger oxidation of C21-ketosteroids from *Rhodococcus rhodochrous*: the isolation and characterization. *Biochim Biophys Acta* 1251:115–124. [https://doi.org/10.1016/0167-4838\(95\)00090-h](https://doi.org/10.1016/0167-4838(95)00090-h)
32. Morii S, Sawamoto S, Yamauchi Y, Miyamoto M, Iwami M, Itagaki E. 1999. Steroid monooxygenase of *Rhodococcus rhodochrous*: sequencing of the genomic DNA, and hyperexpression, purification, and characterization of the recombinant enzyme. *J Biochem* 126:624–631. <https://doi.org/10.1093/oxfordjournals.jbchem.a022494>
33. Franceschini S, van Beek HL, Pennetta A, Martinoli C, Fraaije MW, Mattevi A. 2012. Exploring the structural basis of substrate preferences in Baeyer-Villiger monooxygenases: insight from steroid monooxygenase. *J Biol Chem* 287:22626–22634. <https://doi.org/10.1074/jbc.M112.372177>
34. Van Beilen JB, Mourlane F, Seeger MA, Kovac J, Li Z, Smits THM, Fritsche U, Witholt B. 2003. Cloning of Baeyer-Villiger monooxygenases from *Comamonas*, *Xanthobacter* and *Rhodococcus* using polymerase chain reaction with highly degenerate primers. *Environ Microbiol* 5:174–182. <https://doi.org/10.1046/j.1462-2920.2003.00401.x>
35. Hördt A, López MG, Meier-Kolthoff JP, Schleuning M, Weinhold L-M, Tindall BJ, Gronow S, Kyrpides NC, Woyke T, Göker M. 2020. Analysis of 1,000+ type-strain genomes substantially improves taxonomic classification of *Alphaproteobacteria*. *Front Microbiol* 11:468. <https://doi.org/10.3389/fmicb.2020.00468>
36. Fujii K, Satomi M, Morita N, Motomura T, Tanaka T, Kikuchi S. 2003. *Novosphingobium tardaugens* sp. nov., an oestradiol-degrading bacterium isolated from activated sludge of a sewage treatment plant in Tokyo. *Int J Syst Evol Microbiol* 53:47–52. <https://doi.org/10.1099/ijs.0.02301-0>
37. Ibero J, Sanz D, Galán B, Díaz E, García JL. 2019. High-quality whole-genome sequence of an estradiol-degrading strain, *Novosphingobium tardaugens* NBRC 16725. *Microbiol Resour Announc* 8:e01715-18. <https://doi.org/10.1128/MRA.01715-18>
38. Fürst M, Gran-Scheuch A, Aalbers FS, Fraaije MW. 2019. Baeyer-Villiger monooxygenases: tunable oxidative biocatalysts. *ACS Catal* 9:11207–11241. <https://doi.org/10.1021/acscatal.9b03396>
39. Tolmie C, Smit MS, Opperman DJ. 2019. Native roles of Baeyer-Villiger monooxygenases in the microbial metabolism of natural compounds. *Nat Prod Rep* 36:326–353. <https://doi.org/10.1039/c8np00054a>
40. Mirza IA, Yachnin BJ, Wang S, Grosse S, Bergeron H, Imura A, Iwaki H, Hasegawa Y, Lau PCK, Berghuis AM. 2009. Crystal structures of cyclohexanone monooxygenase reveal complex domain movements and a sliding cofactor. *J Am Chem Soc* 131:8848–8854. <https://doi.org/10.1021/ja9010578>
41. Ji X, Tu J, Song Y, Zhang C, Wang L, Li Q, Ju J. 2020. A luciferase-like monooxygenase and flavin reductase pair AbmE2/AbmZ catalyzes Baeyer-Villiger oxidation in neoabysomicin biosynthesis. *ACS Catal* 10:2591–2595. <https://doi.org/10.1021/acscatal.9b05488>
42. Ramos JL, Martínez-Bueno M, Molina-Henares AJ, Terán W, Watanabe K, Zhang X, Gallegos MT, Brennan R, Tobes R. 2005. The TetR family of transcriptional repressors. *Microbiol Mol Biol Rev* 69:326–356. <https://doi.org/10.1128/MMBR.69.2.326-356.2005>
43. Horinouchi M, Kurita T, Yamamoto T, Hatori E, Hayashi T, Kudo T. 2004. Steroid degradation gene cluster of *Comamonas testosteroni* consisting of 18 putative genes from meta-cleavage enzyme gene *tesB* to regulator gene *tesR*. *Biochem Biophys Res Commun* 324:597–604. <https://doi.org/10.1016/j.bbrc.2004.09.096>
44. Fernández-Cabezón L, Galán B, García JL. 2018. Unravelling a new catabolic pathway of C-19 steroids in *Mycobacterium smegmatis*. *Environ Microbiol* 20:1815–1827. <https://doi.org/10.1111/1462-2920.14114>
45. Sambrook J, Russell DW. 2001. Molecular cloning: a laboratory manual. 3rd ed. CSHL Press, New York, NY, USA.
46. Godoy MS, de Miguel SR, Prieto MA. 2023. A singular PpaA/AerR-like protein in *Rhodospirillum rubrum* rules beyond the boundaries of photosynthesis in response to the intracellular redox state. *mSystems* 8:e00702-23. <https://doi.org/10.1128/mSystems.00702-23>
47. Sharma SB, Signer ER. 1990. Temporal and spatial regulation of the symbiotic genes of *Rhizobium meliloti* in planta revealed by transposon *Tn5-gusA*. *Genes Dev* 4:344–356. <https://doi.org/10.1101/gad.4.3.344>
48. Ibero J, Galán B, García JL. 2021. Identification of the EcdR estrogen-dependent repressor in *Caenibius tardaugens* NBRC 16725: construction of a cellular estradiol biosensor. *Genes (Basel)* 12:1846. <https://doi.org/10.3390/genes12121846>
49. Guevara G, Fernández de las Heras L, Perera J, Navarro Llorens JM. 2017. Functional differentiation of 3-ketosteroid Δ^1 -dehydrogenase isozymes in *Rhodococcus ruber* strain Chol-4. *Microb Cell Fact* 16:42. <https://doi.org/10.1186/s12934-017-0657-1>
50. Li X, Chen T, Peng F, Song S, Yu J, Sidoine DN, Cheng X, Huang Y, He Y, Su Z. 2021. Efficient conversion of phytosterols into 4-androstene-3,17-dione and its C1,2-dehydrogenized and 9 α -hydroxylated derivatives by engineered *Mycobacteria*. *Microb Cell Fact* 20:158. <https://doi.org/10.1186/s12934-021-01653-9>
51. van Oosterwijk N, Knol J, Dijkhuizen L, van der Geize R, Dijkstra BW. 2012. Structure and catalytic mechanism of 3-ketosteroid- δ^4 -(5 α)-dehydrogenase from *Rhodococcus jostii* RHA1 genome. *J Biol Chem* 287:30975–30983. <https://doi.org/10.1074/jbc.M112.374306>
52. García JL, Uhlir I, Galán B. 2012. Catabolism and biotechnological applications of cholesterol degrading bacteria. *Microb Biotechnol* 5:679–699. <https://doi.org/10.1111/j.1751-7915.2012.00331.x>
53. Han S, Liu X, He B, Zhai X, Yuan C, Li Y, Lin W, Wang H, Zhang B. 2024. Efficient production of 9,22-dihydroxy-23,24-bisnorchole-4-ene-3-one from phytosterols by modifying multiple genes in *Mycobacterium fortuitum*. *Int J Mol Sci* 25:3579. <https://doi.org/10.3390/ijms25073579>
54. Hernández-Fernández G, Acedos MG, de la Torre I, Ibero J, García JL, Galán B. 2024. Improving the production of 22-hydroxy-23,24-bisnorchole-4-ene-3-one in *Mycobacterium smegmatis*. *Microb Biotechnol* 17:e14551. <https://doi.org/10.1111/1751-7915.14551>
55. Hernández-Fernández G, Acedos MG, García JL, Galán B. 2024. Identification of the aldolase responsible for the production of 22-hydroxy-23,24-bisnorchole-4-ene-3-one from natural sterols in *Mycobacterium smegmatis*. *Microb Biotechnol* 17:e14270. <https://doi.org/10.1111/1751-7915.14270>
56. Tiwari R, Köffel R, Schneider R. 2007. An acetylation/deacetylation cycle controls the export of sterols and steroids from *S. cerevisiae*. *EMBO J* 26:5109–5119. <https://doi.org/10.1038/sj.emboj.7601924>
57. Paio A, Fogal S, February MR. 2017. New enzymatic process for the preparation of testosterone and esters thereof. US patent US20170029863A1, issued
58. Bisagni S, Hatti-Kaul R, Mamo G. 2014. Cloning, expression and characterization of a versatile Baeyer-Villiger monooxygenase from *Dietzia* sp. D5. *AMB Express* 4:23. <https://doi.org/10.1186/s13568-014-0023-1>
59. Kolek T, Szpineter A, Swizdor A. 2009. Studies on Baeyer-Villiger oxidation of steroids: DHEA and pregnenolone Δ^1 -lactonization pathways in *Penicillium camemberti* AM83. *Steroids* 74:859–862. <https://doi.org/10.1016/j.steroids.2009.05.007>
60. Isupov MN, Schröder E, Gibson RP, Beecher J, Donadio G, Saneei V, Dcunha SA, McGhie EJ, Sayer C, Davenport CF, Lau PCK, Hasegawa Y, Iwaki H, Kadow M, Balke K, Bornscheuer UT, Bourenkov G, Littlechild JA. 2018. The oxygenating constituent of 3,6-diketocamphane monooxygenase from the CAM plasmid of *Pseudomonas putida*: the first crystal structure of a type II Baeyer-Villiger monooxygenase. *Acta Crystallogr D Struct Biol* 74:379. <https://doi.org/10.1107/S205979831800150X>
61. Iwaki H, Grosse S, Bergeron H, Leisch H, Morley K, Hasegawa Y, Lau PCK. 2013. Camphor pathway redux: functional recombinant expression of 2,5- and 3,6-diketocamphane monooxygenases of *Pseudomonas putida*

- ATCC 17453 with their cognate flavin reductase catalyzing Baeyer-Villiger reactions. *Appl Environ Microbiol* 79:3282–3293. <https://doi.org/10.1128/AEM.03958-12>
62. Zhao A, Zhang X, Li Y, Wang Z, Lv Y, Liu J, Alam MA, Xiong W, Xu J. 2021. *Mycobacterium* cell factory for the production of steroid-based drug intermediates. *Biotechnol Adv* 53:107860. <https://doi.org/10.1016/j.biotechadv.2021.107860>
 63. Wójcik P, Głanowski M, Wojtkiewicz AM, Rohman A, Szaleniec M. 2021. Universal capability of 3-ketosteroid Δ^1 -dehydrogenases to catalyze Δ^1 -dehydrogenation of C17-substituted steroids. *Microb Cell Fact* 20:119. <https://doi.org/10.1186/s12934-021-01611-5>
 64. Qin D, Ma C, Hu A, Zhang F, Hu H, Yu C-P. 2016. *Altererythrobacter estronivorus* sp. nov., an estrogen-degrading strain isolated from Yundang lagoon of Xiamen city in China. *Curr Microbiol* 72:634–640. <https://doi.org/10.1007/s00284-016-0995-y>
 65. Hall BW, Kontur WS, Neri JC, Gille DM, Noguera DR, Donohue TJ. 2023. Production of carotenoids from aromatics and pretreated lignocellulosic biomass by *Novosphingobium aromaticivorans*. *Appl Environ Microbiol* 89:e0126823. <https://doi.org/10.1128/aem.01268-23>
 66. Metz F, Olsen AM, Lu F, Myers KS, Allemann MN, Michener JK, Noguera DR, Donohue TJ. 2024. Catabolism of β -5 linked aromatics by *Novosphingobium aromaticivorans*. *mBio* 15:e01718-24. <https://doi.org/10.1128/mbio.01718-24>
 67. Vilbert AC, Kontur WS, Gille D, Noguera DR, Donohue TJ. 2024. Engineering *Novosphingobium aromaticivorans* to produce *cis,cis*-muconic acid from biomass aromatics. *Appl Environ Microbiol* 90:e01660-23. <https://doi.org/10.1128/aem.01660-23>
 68. Horinouchi M, Kurita T, Hayashi T, Kudo T. 2010. Steroid degradation genes in *Comamonas testosteroni* TA441: isolation of genes encoding a $\Delta^4(5)$ -isomerase and 3α - and 3β -dehydrogenases and evidence for a 100 kb steroid degradation gene hot spot. *J Steroid Biochem Mol Biol* 122:253–263. <https://doi.org/10.1016/j.jsbmb.2010.06.002>
 69. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30:2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>
 70. Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nat Methods* 9:357–359. <https://doi.org/10.1038/nmeth.1923>
 71. Anders S, Pyl PT, Huber W. 2015. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics* 31:166–169. <https://doi.org/10.1093/bioinformatics/btu638>
 72. Gu Z, Eils R, Schlesner M. 2016. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics* 32:2847–2849. <https://doi.org/10.1093/bioinformatics/btw313>
 73. Uhía I, Galán B, Morales V, García JL. 2011. Initial step in the catabolism of cholesterol by *Mycobacterium smegmatis* mc²155. *Environ Microbiol* 13:943–959. <https://doi.org/10.1111/j.1462-2920.2010.02398.x>
 74. Herrero M, de Lorenzo V, Timmis KN. 1990. Transposon vectors containing non-antibiotic resistance selection markers for cloning and stable chromosomal insertion of foreign genes in Gram-negative bacteria. *J Bacteriol* 172:6557–6567. <https://doi.org/10.1128/jb.172.11.6557-6567.1990>
 75. Wirth R, Friesenegger A, Fiedler S. 1989. Transformation of various species of Gram-negative bacteria belonging to 11 different genera by electroporation. *Mol Gen Genet* 216:175–177. <https://doi.org/10.1007/BF00332248>
 76. Bradford MM. 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* 72:248–254. [https://doi.org/10.1016/0003-2697\(76\)90527-3](https://doi.org/10.1016/0003-2697(76)90527-3)