



Universidad
de Alcalá

**Departamento de Bioquímica
y Biología Molecular**

**AISLAMIENTO Y CARACTERIZACIÓN DE GENES DE
LA RUTA BIOSINTÉTICA DE UN ANTIFÚNGICO
OXOPENTAENO PRODUCIDO POR *Streptomyces* sp.**

TESIS DOCTORAL

**Aránzazu Sanz López de Alda
Centro Nacional de Biotecnología (CSIC)
Madrid, 2005**

D. FRANCISCO MALPARTIDA ROMERO, Investigador del Consejo Superior de Investigaciones Científicas (CSIC),

CERTIFICA: que la Tesis Doctoral titulada “**Aislamiento y caracterización de genes de la ruta biosintética de un antifúngico oxopentaeno producido por *Streptomyces sp.***”, presentada por D^a. ARÁNZAZU SANZ LÓPEZ DE ALDA para optar al grado de Doctor por la Universidad de Alcalá, ha sido realizada bajo su dirección en el Departamento de Biotecnología Microbiana del Centro Nacional de Biotecnología (CSIC) y reúne los requisitos exigidos para su defensa ante el Tribunal.

Madrid, 17 de octubre de 2005

El Director de la Tesis,

Dr. D. Francisco Malpartida Romero

D. ÁNGEL HERRÁEZ SÁNCHEZ, Profesor Titular de la Universidad de Alcalá,

INFORMA: que el trabajo titulado “**Aislamiento y caracterización de genes de la ruta biosintética de un antifúngico oxopentaeno producido por *Streptomyces* sp.**” ha sido realizado por D^a. ARÁNZAZU SANZ LÓPEZ DE ALDA bajo su tutoría, dentro del programa de Doctorado Biomedicina del Departamento de Bioquímica y Biología Molecular de la Universidad de Alcalá, y, a su juicio, cumple todos los requisitos para su defensa como Tesis Doctoral.

Alcalá de Henares, 17 de octubre de 2005

V^oB^o

La Subdirectora del Departamento
de Bioquímica y Biología Molecular

V^oB^o

El Tutor de Tesis

Prof. Dra. D^a. Pilar Sancho López

Prof. Dr. D. Ángel Herráez Sánchez

*A Luis, a mis padres y hermano,
por su cariño y apoyo*

1. INTRODUCCIÓN

1.1.	<i>Streptomyces</i>	1-3
	- Características generales	
	- Ciclo de vida	
	- Organización genómica	
1.2.	Metabolitos secundarios	3-5
	- Generalidades	
	- Antibióticos	
	- Clasificación de los antibióticos	
1.3.	Poliquétidos	5-14
	- Mecanismo de síntesis de las PKSs	
	- Tipos de PKSs	
	- Clasificación de los poliquétidos	
1.4.	PKSs Tipo I	14-18
	- Organización génica	
	- Modo de acción de las PKSs Tipo I	
1.5.	Manipulación genética de rutas biosintéticas de poliquétidos	19
1.6.	Antecedentes y objetivos del presente trabajo	20

2. MATERIALES Y MÉTODOS

2.1.	Microorganismos utilizados	21
2.2.	Crecimiento y conservación	21-22
2.3.	Medios de cultivo	22-24
2.4.	Selección con antibióticos	24
2.5.	Vectores	25-29
2.6.	Aislamiento de ADN	30
	- Aislamiento de ADN plasmídico de <i>E. coli</i>	
	- Aislamiento de ADN plasmídico de <i>Streptomyces</i>	
	- Aislamiento de ADN total de <i>Streptomyces</i>	
	- Aislamiento de ADN de fagos de <i>E. coli</i>	
	- Aislamiento de ADN de fagos de <i>Streptomyces</i>	
2.7.	Obtención de alto título de fagos	31
	- Obtención de alto título de fagos de <i>E. coli</i>	
	- Obtención de alto título de fagos de <i>Streptomyces</i>	
2.8.	Introducción de ADN	31-32
	- Transformación de <i>E.coli</i>	
	- Transformación de <i>Streptomyces</i>	
	- Transfección de <i>Streptomyces</i>	
	- Infección en <i>Streptomyces</i>	
	- Conjugación entre <i>E. coli</i> y <i>Streptomyces</i>	

2.9.	Manipulación enzimática del ADN	32-33
	- Digestión del ADN	
	- Desfosforilación de extremos cohesivos	
	- Relleno de extremos protuberantes de ADN	
	- Ligación de fragmentos de ADN	
	- Amplificación de ADN por PCR	
2.10.	Electroforesis de ADN	33-34
	- Electroforesis en geles de agarosa	
	- Purificación de fragmentos de ADN a partir de geles de agarosa	
2.11.	Hibridación de ADN	34
	- Transferencia de ADN desde geles de agarosa: <i>Southern Blot</i>	
	- Hibridación de ADN	
2.12.	Marcaje radiactivo de fragmentos de ADN	35
	- Marcaje por desplazamiento de mellas: <i>Nick Translation</i>	
	- Marcaje mediante el relleno de extremos protuberantes	
2.13.	Secuenciación de ADN	35
2.14.	Aislamiento y manipulación de ARN	36
	- Aislamiento de ARN total de <i>Streptomyces</i>	
	- Electroforesis de ARN en geles desnaturalizantes	
2.15.	Ensayos de protección a endonucleasa a S1	36-38
	- Ensayos de protección a endonucleasa a S1 de alta resolución	
	- Marcaje con Polinucleótido Quinasa T4 y [γ ³² P]dATP	
	- Electroforesis en geles de acrilamida-urea 6%	
	- Ensayos de protección a endonucleasa a S1 de baja resolución	
	- Electroforesis en geles alcalinos	
2.16.	Construcción de una genoteca de ADN total de <i>Streptomyces</i> sp. CO-AC25	38
2.17.	Bioensayos frente a <i>Penicillium</i> y <i>Micrococcus</i>	38
2.18.	Análisis por HPLC	39
2.19.	Extracción de metabolitos activos	39-40
	- Fraccionamiento en columna Kromasil C8 de fase reversa	
2.20.	Soporte informático	40

3. RESULTADOS

3.1.	Caracterización de metabolitos activos en <i>Streptomyces</i> sp. CO-AC25	41-44
	- Caracterización de las actividades bactericida y fungicida	
	- Determinación de la estructura química del compuesto fungicida	
3.2.	Análisis de sistemas de manipulación genética	44-46
3.3.	Aislamiento y caracterización de genes biosintéticos de los compuestos bioactivos....	46-57
	- Escrutinio de la genoteca de <i>Streptomyces</i> sp. CO-AC25	
	- Aislamiento de genes biosintéticos a partir de la sonda heteróloga	
	- Estudios de disrupción génica	

-	Análisis de la actividad biocida de los exconjugantes	
-	Ampliación de la región clonada del <i>cluster</i> de flavofungina I	
3.4.	Secuenciación y determinación de regiones abiertas de lectura de la PKS de flavofungina I	57-59
3.5.	Funciones deducidas de las regiones abiertas de lectura de la PKS de flavofungina I	59-65
3.6.	Análisis transcripcional de los genes <i>flvA</i> y <i>flvB</i>	66-69
-	Análisis transcripcional del gen <i>flvA</i>	
-	Análisis transcripcional del gen <i>flvB</i>	
3.7.	Generación de un módulo de arranque híbrido en la proteína FlvA	69-77
-	Construcción del gen <i>flvA</i> híbrido (<i>flvA*</i>)	
-	Generación de FlvA* completa mediante recombinación “ <i>in vivo</i> ”	
-	Evaluación de la actividad biológica	
3.8.	Funciones deducidas de las regiones abiertas de lectura de la PKS del compuesto bactericida	78
3.9.	Caracterización de las regiones abiertas de lectura de la PKS del compuesto bactericida	79-82
3.10.	Expresión de los genes <i>orf8</i> y <i>orf9</i> en <i>Streptomyces lividans</i> TK21	82-87
-	Estudio de la expresión de los genes <i>orf8</i> y <i>orf9</i> en <i>S. lividans</i> TK21::ABC y en <i>S. lividans</i> TK21::ABCery	
4.	DISCUSIÓN	88-95
5.	CONCLUSIONES	96
6.	ANEXOS	
-	Anexo I	97-130
-	Anexo II.....	131-144
-	Anexo III.....	145-147
7.	BIBLIOGRAFÍA	148-159

Abreviaturas

aa	aminoácido
ACP	proteína transportadora de grupos acilo
ADN	ácido desoxirribonucleico
<i>amp^r</i>	gen de resistencia a ampicilina
<i>ap^r</i>	gen de resistencia a apramicina
ARN	ácido ribonucleico
ARO	aromatasa
AT	aciltransferasa
ATP	adenosina 5'-trifosfato
CHS	sintasa chalcona
CIAP	fosfatasa alcalina de intestino de ternera
CLF	<i>Chain Length Factor</i>
<i>cm^r</i>	gen de resistencia a cloranfenicol
CoA	Coenzima A
cpm	cuentas por minuto
CYC	ciclasa
dATP	desoxiadenosina 5'-trifosfato
dCTP	desoxicitidina 5'-trifosfato
dGTP	desoxiguanosina 5'-trifosfato
DH	deshidratasa
DMSO	dimetil-sulfóxido
EDTA	ácido etilendiamino tetracético
ER	enoilreductasa
FAS	sintasa de ácidos grasos
β-gal	β-galactosidasa
<i>hyg^r</i>	gen de resistencia a higromicina
HPLC	<i>High Performance Liquid Chromatography</i>
IPTG	isopropil tio β-D-galactósido
kb	kilobase

kDa	kilodalton
Klenow	fragmento mayor de la ADN polimerasa de <i>E. coli</i>
<i>km^r</i>	gen de resistencia a kanamicina
KR	ketorreductasa
KS	ketosintasa
Mb	megabase
mM	milimolar
MOPS	ácido 3N-(morfolino)-propano-sulfónico
<i>neo^r</i>	gen de resistencia a neomicina
nm	nanómetros
pb	pares de bases
PCR	reacción en cadena de la polimerasa
PKS	poliquétido sintasa
PNK	polinucleótido Quinasa
RMN	resonancia magnética nuclear
rpm	revoluciones por minuto
TBE	tris-ácido-bórico-EDTA
Tc^r	gen de resistencia a tetraciclina
TE	tioesterasa
TEA	trietanolamina
Tris	tris-2-amino-2 (hidroximetil)-1,3-propanodiol
<i>tsr^r</i>	gen de resistencia a triostrepton
RBS	sitio de unión a ribosoma
ufp	unidades formadoras de placa
UV	ultravioleta
X-Gal	5-bromo-4-cloro-3-indolil-β-D-galactopiranosido

INTRODUCCIÓN



1.1. *Streptomyces*

Características generales

El género *Streptomyces* pertenece a la familia Streptomycetaceae y al orden de los Actinomycetales. Engloba bacterias filamentosas, miceliales, Gram-positivas, aerobias, capaces de utilizar un gran número de compuestos orgánicos como fuente de carbono y energía y con un elevado contenido en guaninas y citosinas (74% G+C) en su genoma (Woese, 1987). Están ampliamente distribuidos en la naturaleza, siendo el suelo su hábitat más común.

Se parecen a los hongos por ser pleomórficos y producir un micelio normalmente diferenciado en micelio vegetativo y micelio aéreo, pero presentan características típicas de procariontes tales como la carencia de núcleo, mitocondrias y cloroplastos (Stackebrandt *et al.*, 1992).

La importancia de estos organismos radica en su capacidad de producir una gran diversidad de metabolitos secundarios, la mayoría de ellos con actividad biológica. Dentro de estos metabolitos secundarios se encuentran desde análogos de metabolitos primarios (aminoácidos, nucleótidos o polipéptidos) hasta antibióticos. Las bacterias del género *Streptomyces* son productoras de más de la mitad de los antibióticos conocidos. De los 12.000 antibióticos caracterizados en 1995: un 55% eran producidos por actinomicetos del género *Streptomyces*, un 11% por otros actinomicetos, un 12% por bacterias no filamentosas y el 22% restante por hongos filamentosos (Berdy, 1995).

Las bacterias del género *Streptomyces* también son productoras de gran cantidad de enzimas extracelulares de gran interés en el sector industrial, entre las que destacan: proteasas (Henderson *et al.*, 1987), celulasas, nucleasas, amilasas (Long *et al.*, 1987), lipasas (Jiménez *et al.*, 1987), quitinasas y xilanasas (Morosoli *et al.*, 1986).

Ciclo de vida

Los estreptomicetos presentan un ciclo de vida complejo que implica procesos de diferenciación morfológica y fisiológica. Estas bacterias son capaces de colonizar sustratos relativamente secos con restos de materia orgánica, formando una red de hifas tabicadas ramificadas que dan lugar al micelio sustrato. Estas hifas obtienen los nutrientes de la degradación del material orgánico insoluble gracias a numerosas enzimas hidrolíticas (Chater, 1984). En una primera etapa, las zonas más alejadas de la fuente de nutrientes empiezan a acumular sustancias de reserva (lípidos, glucógeno,...), hasta que en un determinado momento y debido a la carencia de nutrientes, se reciben una serie de señales justamente en esta zona que disparan la expresión de genes implicados en la formación del micelio aéreo. Se produce así, el desarrollo de hifas que emergen del micelio sustrato para dar lugar al micelio aéreo. Estas hifas se van a nutrir de los productos de degradación del micelio sustrato, y en una segunda etapa van a sufrir un proceso de curvatura, enrollamiento, formación de septos y engrosamiento de la pared celular para dar lugar a una cadena de esporas uninucleares, que se liberarán al medio y que con las condiciones adecuadas, germinarán y desarrollarán un nuevo micelio sustrato (Figura 1).

La producción de metabolitos secundarios, por lo general, coincide con el inicio de la diferenciación morfológica, una vez terminado el desarrollo vegetativo (Champness y Chater, 1994).

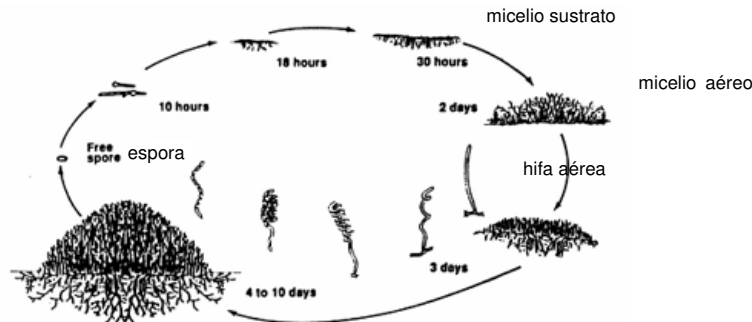


Fig. 1. Ciclo de vida de *Streptomyces*.

Organización genómica

Mediante el análisis de fragmentos de restricción, en electroforesis de campo pulsante, de cromosomas de diferentes especies de *Streptomyces* como *S. coelicolor* (Kieser *et al.*, 1992), *S. ambofaciens* (Leblond *et al.*, 1990), *S. lividans* (Leblond *et al.*, 1993), *S. griseus* (Lezhava *et al.*, 1995) o *S. rimosus* (Pandza *et al.*, 1997), se ha llegado a demostrar la linealidad de su genoma (Lin *et al.*, 1993) así como su tamaño aproximado de 8 Mb. Se piensa, que su gran tamaño pueda deberse a la presencia de gran número de regiones no codificantes (Robinson *et al.*, 1981), repeticiones de ADN en tándem (Shrempf, 1985) o duplicación de genes con la misma función aparente (Ohnuki *et al.*, 1985; Tohyama *et al.*, 1987). Se han detectado la presencia de secuencias repetidas invertidas en los extremos de los cromosomas que tienden a ser muy inestables y pueden inducir deleciones de hasta 2 Mb. Se postula que estas regiones tengan la función de evitar la deleción de otras regiones del cromosoma que sean imprescindibles (Redenbach *et al.*, 1996).

En la actualidad se encuentran secuenciados en su totalidad el genoma de *Streptomyces coelicolor* A3(2) (Bentley *et al.*, 2002) y el de *Streptomyces avermitilis* (Omura *et al.*, 2001; Ikeda *et al.*, 2003) y está próximo a concluir el de *S. clavuligerus*, lo que ha aportado un gran avance en el conocimiento de la organización génica en *Streptomyces*.

El genoma de *Streptomyces coelicolor* es un cromosoma lineal de 8.667.507 pb con la presencia de 7.825 probables genes, dentro de los cuales se incluyen más de 20 *clusters* implicados en la síntesis de metabolitos secundarios. El genoma de *Streptomyces avermitilis* también es un cromosoma lineal, en este caso su tamaño es de 9.025.608 pb que codifican al menos 7.574 regiones abiertas de lectura, dentro de las cuales se han identificado 30 *clusters* relacionados con el metabolismo secundario, aproximadamente un 6,6% de todo el genoma.

La distribución de los genes en el genoma no es aleatoria, se ha visto que los genes que resultan esenciales (genes implicados en la división celular, en la replicación del ADN, transcripción, traducción y biosíntesis de aminoácidos) se encuentran en el centro del cromosoma, mientras que los genes codificantes para funciones tales como la síntesis de

metabolitos secundarios o enzimas hidrolíticas, se localizan en los brazos del cromosoma. Se piensa que se evita así la pérdida de genes esenciales a causa de deleciones que puedan ocurrir en los extremos de los brazos.

Los elementos extracromosomales son frecuentes en *Streptomyces*. Se han encontrado plásmidos circulares cerrados covalentemente de tamaño y número de copias muy variable (Hopwood *et al.* 1986) y plásmidos lineales con proteínas asociadas a sus extremos (Kinashi y Shimaji-Murayama, 1987). También está descrita la existencia de plásmidos capaces de integrarse en sitios específicos del cromosoma de *Streptomyces* (Brown *et al.*, 1988; Mazodier *et al.*, 1990; Combes *et al.*, 2002). Además de plásmidos existen fagos de *Streptomyces*, todos ellos de doble cadena de ADN y que pueden ser tanto virulentos como atemperados (Lomovskaya *et al.*, 1980). El más conocido es el fago atemperado ϕ C31 de 41,4 kb, a partir del cual, mediante una serie de manipulaciones como pueden ser inserciones, deleciones o recombinaciones, se han obtenido numerosos vectores fágicos (Hopwood *et al.* 1985). Tanto plásmidos como fagos sirven como herramientas para el abordaje de estudios genéticos y experimentos de biosíntesis combinatorial en *Streptomyces*.

Dentro del cromosoma de estas bacterias también se han encontrado fragmentos que corresponden a elementos genéticos transponibles.

1.2. METABOLITOS SECUNDARIOS

Generalidades

El término metabolito secundario hace referencia a aquellos compuestos de gran diversidad estructural que no son esenciales para el crecimiento y reproducción del organismo productor.

Entre las características comunes que presentan estos metabolitos secundarios destacan (Vining, 1992):

- La especificidad de la cepa productora.
- La variabilidad de estructuras químicas y actividades biológicas.
- Su síntesis en la etapa final del crecimiento exponencial y durante la fase estacionaria, a partir de intermediarios del metabolismo primario y mediante rutas biosintéticas largas y complejas.
- El control fisiológico de su producción como respuesta a factores ambientales.
- La agrupación en el cromosoma de los genes implicados en la biosíntesis de dichos metabolitos en forma de *clusters* (Martín y Liras, 1989), asociados a genes reguladores de su expresión.

Antibióticos

Los antibióticos son metabolitos secundarios de bajo peso molecular, cuya actividad biológica se orienta a inhibir el crecimiento de microorganismos distintos a los productores a bajas concentraciones, confiriendo una ventaja selectiva para éstos.

Además de la actividad antibiótica, los metabolitos secundarios pueden presentar otras actividades de interés. Así existen metabolitos con actividad inmunosupresora, como el macrólido FK506 (Kino *et al.* 1987); metabolitos capaces de inducir la motilidad gástrica, como la eritromicina (Itoh *et al.* 1984); y metabolitos secundarios que presentan propiedades antihipertensivas o antihipercolesterolémicas.

Junto a los *clusters* de biosíntesis de los antibióticos se encuentran también los genes implicados en la resistencia frente a los mismos en los organismos productores. Estos genes de resistencia actúan con diversos mecanismos:

- Modificación de la estructura del antibiótico en el interior celular bien mediante N-acetilación de grupos amino por acción de acetiltransferasas, como por O- fosforilación de grupos hidroxilo por acción de fosfotransferasas. Así, el antibiótico modificado pierde su actividad biológica (Cundliffe, 1989).

- Modificación de las dianas de acción, ya sean ribosomales (ARN ribosomal) o no (enzimas claves en la síntesis de ácidos nucleicos, proteínas o ácidos grasos) (Cundliffe, 1978 y 1989).

- Modificación de la permeabilidad mediante bombas asociadas a la membrana celular, que provocan una salida del antibiótico al exterior celular y pueden impedir su entrada de nuevo. Es el caso de la oxitetraciclina producida por *Streptomyces rimosus* (Ohnuki *et al.*, 1985) que es excretada al exterior celular mediante un mecanismo dependiente de energía.

De entre los mecanismos encargados de secretar antibióticos al exterior celular, cabe destacar la familia de los ABC-transportadores: proteínas que utilizando como fuente de energía la hidrólisis de ATP (de ahí que también se las denomine *ATP-binding cassette*), excretan al medio extracelular diferentes moléculas (azúcares, aminoácidos, oligopéptidos, drogas, etc.) mediante un proceso en contra de gradiente de concentración. Así se evita la acumulación de sustancias tóxicas en el interior celular evitando las consecuencias que dicho acúmulo lleva asociado.

Las proteínas pertenecientes a la familia ABC-transportadores, normalmente están formadas por dos componentes. Un componente hidrofílico, responsable de la unión al ATP y de su posterior hidrólisis y un componente hidrofóbico de interacción con la membrana citoplasmática. En función de cómo se organicen los genes que codifican para estos sistemas, se pueden distinguir tres tipos de ABC-transportadores (Méndez y Salas, 1998): ABC-transportadores Tipo I, ABC-transportadores Tipo II y ABC-transportadores Tipo III. Los primeros están constituidos por dos genes, uno de ellos codifica el componente hidrofílico del sistema y el otro codifica el componente hidrofóbico. En el caso de los ABC-transportadores Tipo II sólo existe un gen responsable de codificar el componente hidrofílico y no existe un componente de unión a membrana citoplasmática. Por último los ABC-transportadores Tipo III,

igual que en el caso anterior solamente poseen un gen que es responsable de codificar los dos componentes del sistema, hidrofílico e hidrofóbico.

Clasificación de los antibióticos

En base a su estructura química los antibióticos se clasifican en :

- **Peptídicos:** se trata de cadenas de aminoácidos sintetizadas en su mayoría mediante péptido sintetasas independientes de ribosomas, que sufrirán después una serie de modificaciones postraduccionales y procesamientos proteolíticos para dar lugar a estructuras muy diversas. En este grupo se encuentran los antibióticos β -lactámicos (penicilinas, cefalosporinas, cefamicinas,...) cuya molécula precursora es el tripéptido L- α -aminoadipoil-L-cisteinil-D-valina (ACV) producida por *Penicillium chrysogenum* y *Streptomyces clavuligerus*.

- **Aminoglucósidos:** son aminoazúcares unidos a un aminociclitol, cabe destacar en este grupo a los antibióticos neomicina, vancomicina, kanamicina, estreptomina, estreptotricina, gentamicina, higromicina,...

- **Poliqúétidos:** son compuestos con estructuras químicas muy variadas, cuya síntesis deriva de una serie de condensaciones sucesivas de metabolitos esenciales catalizadas por una actividad enzimática conocida como poliqúétido sintasas (PKSs). En este grupo se engloban compuestos como la eritromicina, la actinorrodina, la anfotericina B, la oleandomicina, etc.

- **Nucleósidos:** son compuestos que por su estructura actúan como análogos de los nucleósidos que componen los ácidos nucleicos. De entre los nucleósidos producidos por actinomicetos destaca la puromicina.

Existen otros compuestos con actividad antibiótica, que por su estructura no pueden englobarse dentro de los grupos anteriores, es el caso del cloranfenicol y la fosfomicina entre otros.

1.3. POLIQUÉTIDOS

Los poliqúétidos son metabolitos secundarios, producidos mayoritariamente por actinomicetos, que engloban gran número de compuestos con estructuras químicas diversas aunque comparten un patrón común de biosíntesis. Las enzimas encargadas de la síntesis de estos compuestos se conocen como poliqúétido sintasas (PKSs): mediante la adición sucesiva de unidades de ácido carboxílico se construye una cadena carbonada. Las unidades de construcción son de varias clases: acetato, propionato o butirato, aunque a veces pueden ser más complejas. Cada unidad contribuye con dos átomos de carbono al montaje de la cadena

lineal, llevando el carbono- β siempre un grupo ceto; algunos de estos grupos son reducidos a hidroxilos y durante la condensación, algunos, son eliminados por decarboxilación. El hecho de que existan grupos ceto alternos da lugar al nombre de poliquétidos para esta familia de compuestos. Formada la cadena policetónica, ésta sufre ciclación y modificaciones de “adorno” o post- PKS.

El mecanismo de acción de estas enzimas es muy similar al de las sintasas de ácidos grasos de cadena larga.

La cadena policetónica presenta un número de átomos de carbono muy variable, desde 6 como ocurre con el compuesto 6- metilsalicílico, hasta 50 como ocurre con la brevitoxina B. En función de la complejidad de las estructuras químicas, los poliquétidos se clasifican en: aromáticos, macrólidos y poliéteres.

Mecanismo de síntesis de las PKSs

Todos los compuestos poliquétidos tienen en común una ruta biosintética en la que a partir de una unidad iniciadora y por sucesivas incorporaciones de unidades de elongación se genera una cadena policetónica lineal y parcialmente oxidada que posteriormente se ciclará y sufrirá una serie de modificaciones post- PKS para dar lugar a la estructura característica de cada poliquétido.

La enzima clave en la biosíntesis de estos compuestos es una poliquétido sintasa (PKS). Se trata de una actividad muy similar a la de las sintasas de ácidos grasos de cadena larga (FASs), donde intervienen diferentes actividades como: aciltransferasa (AT), ketosintasa (KS), proteína transportadora de grupos acilo (ACP), ketorreductasa (KR), deshidratasa (DH) y enoilreductasa (ER) (Hopwood y Sherman, 1990).

El esquema básico para la formación de los ácidos grasos de cadena larga en animales es uno de los mejor estudiados y sucede mediante la incorporación sucesiva de unidades de elongación a partir de una unidad iniciadora de acetato, ambas unidades en forma de tioésteres CoA. La molécula de acetato, unidad iniciadora, es transferida inicialmente por la acción de una AT desde el CoA al grupo tiol del brazo de 4'-fosfopantoteína del dominio ACP para su posterior condensación. Por su parte, la unidad elongadora, que tras activarse en forma malonil-CoA, es transferida desde el CoA al grupo tiol del brazo de 4'-fosfopantoteína de la proteína transportadora de grupos acilo (ACP) mediante la misma AT que transfiere la unidad iniciadora.

A continuación, tiene lugar la condensación entre ambas moléculas por el dominio KS, condensación que se produce con la consecuente pérdida de un carbono por decarboxilación de la unidad elongadora. El resultado es la formación de un residuo policetónico de 4 carbonos unido al brazo de 4'-fosfopantoteína del dominio ACP, que sufrirá reacciones de β -ketorreducción, deshidratación y enoilreducción. Mediante sucesivas condensaciones de unidades de elongación llega a constituirse la cadena policetónica (Figura 2).

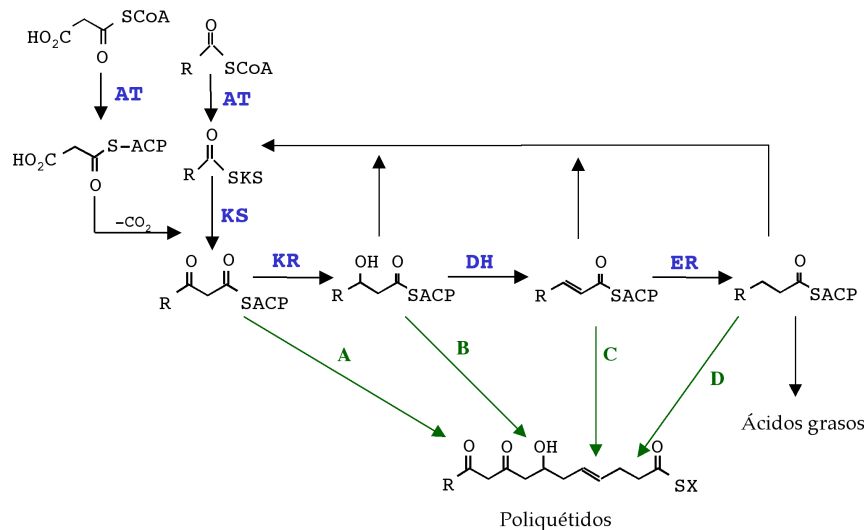


Fig. 2. Ruta básica de biosíntesis de ácidos grasos y poliquétidos. Las letras A, B, C y D se corresponden con los distintos grados de reducción que puede sufrir la cadena policetónica.

Ha de señalarse que, aunque el mecanismo biosintético de los ácidos grasos de cadena larga y los poliquétidos sea similar, existen algunas diferencias (Katz y Donadio, 1993). La unidad iniciadora en ambos casos puede ser acetil-CoA u otro tioéster-CoA, y mientras que la unidad elongadora suele ser malonil-CoA para ácidos grasos y poliquétidos aromáticos puede variar en el caso de los poliquétidos complejos, donde se puede incorporar metilmalonil-CoA o etilmalonil-CoA, originándose cadenas laterales metilo o etilo en la estructura del poliquétido. Otra diferencia es, que en el caso de los poliquétidos, el procesamiento de cada grupo ceto incorporado no es constante, pudiendo dar lugar a la formación de grupos ceto, hidroxilo, enoilo o enlaces C-C totalmente reducidos. Además, una vez sintetizada la cadena policetónica, ésta sufre procesos de plegamiento y ciclación que dan lugar a poliquétidos aromáticos o macrocíclicos y que pueden ir acompañados de glicosilaciones, amidaciones,...

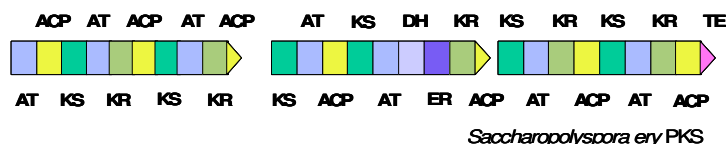
Tipos de PKSs

Se postula que las PKSs pudieran tener un origen evolutivo común a partir de las FASs, y al igual que se habla de FASs Tipo I (características de hongos y vertebrados) y FASs Tipo II (características de bacterias y plantas), por analogía con las mismas, las PKSs también se clasifican en PKSs Tipo I y Tipo II (Hutchinson y Fujii, 1995). Recientemente, se ha establecido el Tipo III como sistema intermedio (Izumikawa *et al.*, 2003).

Las PKSs Tipo I se componen de varias proteínas multifuncionales, cada una de las cuales presenta diferentes dominios catalíticos colocados en el orden preciso en que van a ser utilizados para la elongación de la cadena policetónica (Cortés *et al.*, 1990; Donadio y Hutchinson, 1991; Aparicio *et al.*, 1996). Las PKSs Tipo II se corresponden con un complejo multienzimático formado por varias proteínas que van a actuar de forma reiterada para originar la cadena policetónica. Por último las PKSs Tipo III son un grupo de proteínas homodiméricas algo distintas a las PKSs Tipo I y Tipo II, dentro de las que se engloba la superfamilia de las sintasas chalconas (CHSs) (Moore and Hopke, 2001; Saxena *et al.*, 2003).

Además de la diferencia estructural que existe entre las PKSs Tipo I y Tipo II, el grado de complejidad genético entre ellas también es distinto. Mientras que las PKSs Tipo II se encuentran codificadas en pocos miles de nucleótidos en el genoma, las PKSs Tipo I son extremadamente complejas, llegando a ocupar varias decenas de miles de nucleótidos con unidades secuenciales repetidas en el gen que dan lugar a una organización conocida como modular (Figura 3).

PKS Tipo I



PKS Tipo II

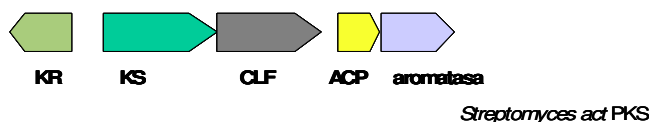


Fig. 3. Esquema de organización de PKSs Tipo I y II.

Clasificación de los poliquétidos

Los poliquétidos son una familia de antibióticos que engloba compuestos con estructuras químicas muy diversas en base a lo cual se pueden clasificar en:

- **Poliquétidos aromáticos:** se trata de compuestos poliquétidos que presentan un anillo aromático en su estructura química. Son compuestos con estructuras químicas muy diversas, pero todos ellos están sintetizados por PKSs Tipo II.

El poliquétido aromático cuya ruta de biosíntesis se encuentra mejor caracterizada es la actinorrodina, producido por *S. coelicolor* A3(2) (Wright y Hopwood, 1976). Se trata del primer *cluster* biosintético de una PKS Tipo II caracterizado genéticamente en detalle, lo que le ha constituido como un modelo para clonar y caracterizar otros *clusters* biosintéticos de PKSs Tipo II.

Todas las PKSs Tipo II implicadas en la biosíntesis de poliquétidos aromáticos tienen una organización muy similar (Figura 4). En todas ellas están presentes tres actividades que en su conjunto constituyen la PKS mínima: KS_{α} , KS_{β} y ACP. La proteína KS_{α} está implicada en la condensación y presenta los centros activos para las actividades KS y AT (Fernández-Moreno *et al.*, 1992). La proteína KS_{β} es homóloga a la KS_{α} pero carece de los sitios activos para las actividades KS y AT y parece ser determinante de la longitud de la cadena policetónica (McDaniel *et al.*, 1993a y b), por lo que se le ha dado el nombre de Factor determinante de la elongación o CLF (*Chain Length Factor*); es necesaria para la biosíntesis de la cadena policetónica presentando una actividad decarboxilasa (Bisang *et al.*, 1999). La tercera proteína que forma parte de la PKS mínima es la proteína transportadora de grupos acilo (ACP). La actividad AT transfiere las diferentes unidades implicadas en el ensamblaje de la cadena

policetónica al grupo tiol del brazo de 4'-fosfopantoteína del ACP, y éste se encarga de aproximarlas al centro activo de la KS, para la condensación.

La biosíntesis de la mayoría de los poliquétidos aromáticos comienza con acetato como unidad iniciadora, salvo en el caso de oxitetraciclina que comienza con malonamida y de daunomicina que comienza con propionato.

Otras actividades enzimáticas implicadas en la formación de los poliquétidos aromáticos son la ya conocida ketorreductasa (KR), que cataliza la ketorreducción de la cadena policetónica creciente; la aromatasas (ARO), que cataliza la aromatización de la cadena naciente y la ciclasa (CYC), implicada en la correcta ciclación de la cadena policetónica. Estas últimas actividades, ARO y CYC, suelen aparecer juntas en una misma proteína.

La PKS Tipo II se emplea en todos los pasos de la síntesis de modo reiterativo, en un proceso en el que los grupos β -carbonilo formados se dejan en su mayoría sin reducir.

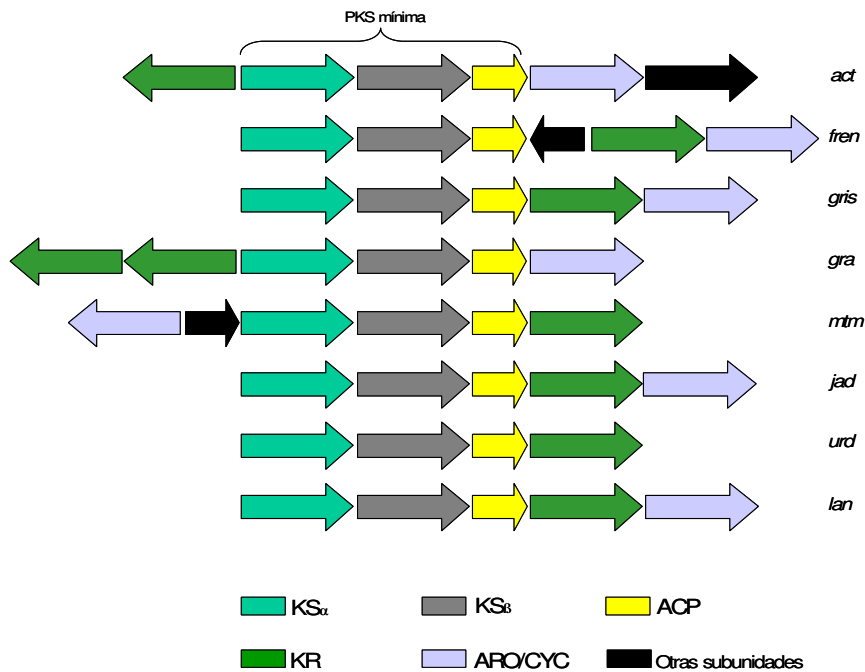


Fig. 4. Representación de los *clusters* de biosíntesis de algunos poliquétidos aromáticos.

Grupo	Producto	Microorganismo productor	Genes de la PKS
Antraquinonas	Actinorrodina	<i>S. coelicolor</i>	<i>act</i>
	Frenolicina/ Nanaomicina	<i>S. roseofulvus</i>	<i>fren</i>
	Griseusina	<i>S. griseus</i>	<i>gris</i>
	Granaticina	<i>S. violaceoruber</i>	<i>gra</i>
	Antraciclinas	Daunomicina	<i>S. sp C5</i>
	Daunorrubicina	<i>S. peucetius</i>	<i>dnr/dps</i>
Similares a Antraciclinas	Tetracenomicina	<i>S. glaucescens</i>	<i>tcm</i>
Tetraciclinas	Oxitetraciclina	<i>S. rimosus</i>	<i>otc</i>
Ácidos aureólicos	Mitramicina	<i>S argillaceus</i>	<i>mtm</i>
Anguciclinas	Jadomicina	<i>S. venezuelae</i>	<i>jad</i>
	Urdamicina	<i>S. fradie</i>	<i>urd</i>
	Landomicina	<i>S. cyanogenus</i>	<i>lan</i>

Tabla 1. Poliquétidos aromáticos cuyas PKSs se encuentran clonadas y caracterizadas.

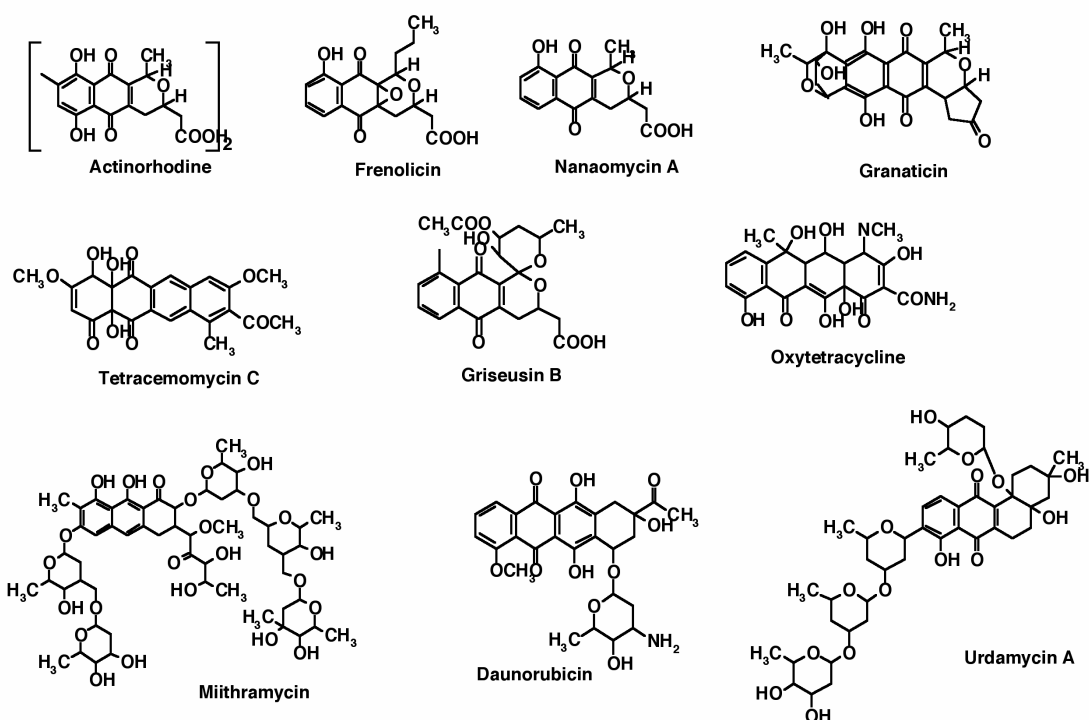


Fig. 5. Estructuras químicas de poliquétidos aromáticos.

- **Macrólidos:** se caracterizan por presentar en su estructura un anillo macrocíclico cerrado por lactonización originado al ciclarse la cadena policetónica correspondiente (Omura y Tanaka, 1984). Son producidos generalmente por bacterias del género *Streptomyces* y otras bacterias filamentosas relacionadas, estando siempre implicadas en su producción PKSs Tipo I. Los macrólidos se pueden clasificar en:

- **macrólidos no polienos:** en los cuales el anillo macrocíclico carece de dobles enlaces conjugados. El primer macrólido no polieno caracterizado fue la picromicina (Omura y Tanaka, 1984), desde entonces son muchos los macrólidos no polienos caracterizados. Entre ellos destaca la eritromicina producida por *Saccharopolyspora erythraea* (Donadio y Katz, 1992) y que actualmente se emplea en clínica por ser un antibiótico de amplio espectro.

Todos ellos presentan un número variable de átomos de carbono que oscila entre 12 y 18, aunque se ha descrito la presencia de 34 átomos de carbono en el caso de las axenomicinas (Martín y McDaniel, 1977). Se caracterizan por presentar en el anillo macrocíclico grupos metilo e hidroxilo y al menos un residuo glucosídico que puede ser crítico para la actividad antibiótica, la cual se fundamenta en la unión al sitio P de la subunidad 50s del ribosoma bacteriano. Su espectro de actividad se restringe a bacterias Gram-positivas, siendo muy limitado frente a bacterias Gram- negativas; frente a eucariotas tienen una actividad muy baja o carecen de ella. Presentan una baja toxicidad por lo que tienen aplicaciones prácticas tanto en medicina como en veterinaria. Su efecto puede ser bacteriostático o bactericida, dependiendo de la especie bacteriana sobre la que actúen, del tamaño del inóculo, de la fase de crecimiento en que se encuentren las bacterias y de la concentración que alcance el antibiótico en el lugar de la infección (Mediavilla, 1997).

Macrólido no polieno	Microorganismo productor	Genes de la PKS
Eritromicina	<i>Sac. erythraea</i>	<i>ery</i>
Tilosina	<i>S. fradiae</i>	<i>tyl</i>
Espiramicina	<i>s. ambofaciens</i>	<i>srm</i>
Oleandomicina	<i>S. antibioticus</i>	<i>ole</i>
Avermectina	<i>S. avermitilis</i>	<i>avr</i>
Carbomicina	<i>S. thermotolerans</i>	<i>car</i>

Tabla 2. Poliquétidos macrólidos no polienos.

Introducción

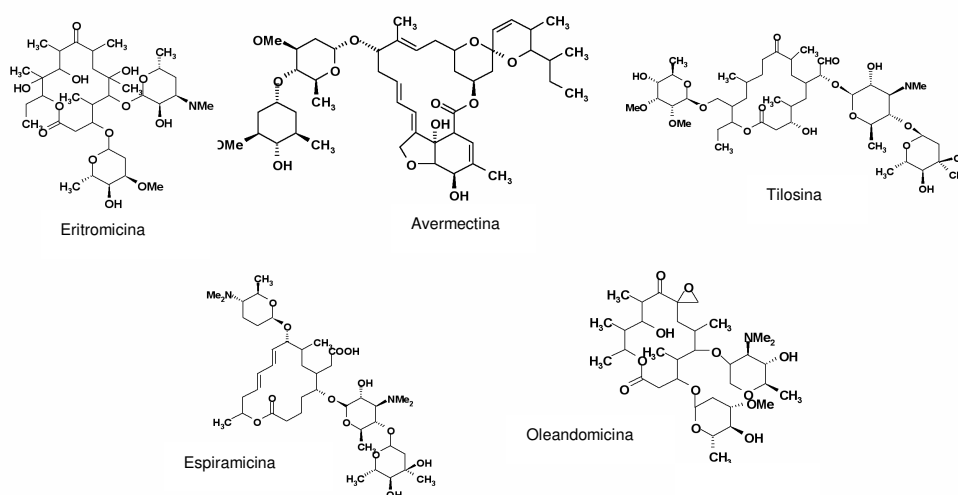


Fig. 6. Estructuras químicas de poliquétidos macrólidos no polienos.

- **macrólidos polienos:** se caracterizan por la presencia de dobles enlaces conjugados en el anillo macrocíclico. Desde el descubrimiento de la nistatina (Hazen y Brown 1950), se han descrito más de 200 antibióticos macrólidos polienos. El número de dobles enlaces conjugados determina el espectro ultravioleta característico de cada uno de ellos. Así se denominan trienos, tetraenos, pentaenos, hexaenos y heptaenos según presenten 3, 4, 5, 6 o 7 dobles enlaces conjugados. La presencia de estos dobles enlaces conjugados determina que los macrólidos polienos sean moléculas anfipáticas, con una región lipofílica plana y una región hidrofílica flexible. El cromóforo, constituido por estos dobles enlaces, es además el responsable de las características físicas y propiedades químicas de los polienos, tales como fuerte absorción de la luz, fotolabilidad y baja solubilidad en agua.

El número de átomos de carbono que constituyen el anillo macrocíclico varía de 20 a 44 átomos, superior al número de átomos de carbono del anillo macrocíclico de los macrólidos no polienos. Pueden presentar un azúcar unido al anillo macrocíclico por un enlace glicosídico a un carbono con un grupo hidroxilo adyacente; suele tratarse de un aminoazúcar, generalmente micosamina (3-amino-3,6-dideoxi-D-manosa) o perosamina (4-amino-4,6-dideoxi-D-manosa). Se han encontrado polienos con azúcares neutros, posibles intermediarios en la biosíntesis de análogos amidados (Gil y Martín, 1997) e incluso polienos carentes del azúcar, por lo éste parece no ser indispensable para la actividad fungicida, aunque favorece la interacción de los polienos con las membranas de las células eucariotas (Hamilton-Miller, 1973).

La actividad biocida de los macrólidos polienos suele ser fungicida; estos compuestos se unen con cierta especificidad al ergosterol presente en la membrana de los hongos, haciéndola permeable y alterando el movimiento de iones a través de ella, lo que afecta a los potenciales de membrana, y conduce finalmente a la muerte celular. Se ha demostrado que los macrólidos polienos también pueden unirse a las membranas con colesterol (Bolard, 1986).

La mayoría de los macrólidos polienos descritos presentan actividad antifúngica y por ello son de gran aplicación en el tratamiento clínico de enfermedades infecciosas causadas por hongos tanto en hombres como animales. El heptaeno anfotericina B (Caffrey *et al.*, 2001) es el principal agente antifúngico utilizado a nivel clínico, y el tetraeno nistatina también tiene una relevancia importante, aunque su uso es mayoritariamente tópico. La pimarcina es un tetraeno de gran uso en el sector industrial, pues se utiliza como conservante alimentario (aditivo E235), preservando la contaminación de los alimentos por micotoxinas.

Estos macrólidos polienos presentan un grado variable de toxicidad debido a que también pueden interaccionar con el colesterol de las membranas de los mamíferos ocasionando problemas hemolíticos y nefrolíticos.

Se conocen otras acciones asociadas a polienos, tales como antivirales (Kessler *et al.*, 1981) o inmunosupresoras. Así, en 1989, se descubrió un antibiótico macrólido pentaeno denominado faerifungina, producido por *S. griseus* var. *autotrophicus*, que además de ser fungicida resultó ser un potente bactericida frente a todas las especies de bacterias Gram-positivas analizadas así como frente a nemátodos y larvas de mosquito (Nair *et al.*, 1989).

Macrólido polieno	Nº de dobles enlaces conjugados	Microorganismo productor	Genes de la PKS
Nistatina	4 (tetraeno)	<i>S. noursei</i>	<i>nys</i>
Pimarcina	4 (tetraeno)	<i>S. natalensis</i>	<i>pim</i>
Rimocidina y CE-108	4 (tetraeno)	<i>S. diastaticus</i> var. 108	<i>rim</i>
Anfotericina B	7 (heptaeno)	<i>S. nodosus</i>	<i>amp</i>
Candidina	7 (heptaeno)	<i>S. griseus</i>	<i>can</i>

Tabla 3. Poliquétidos macrólidos polienos.

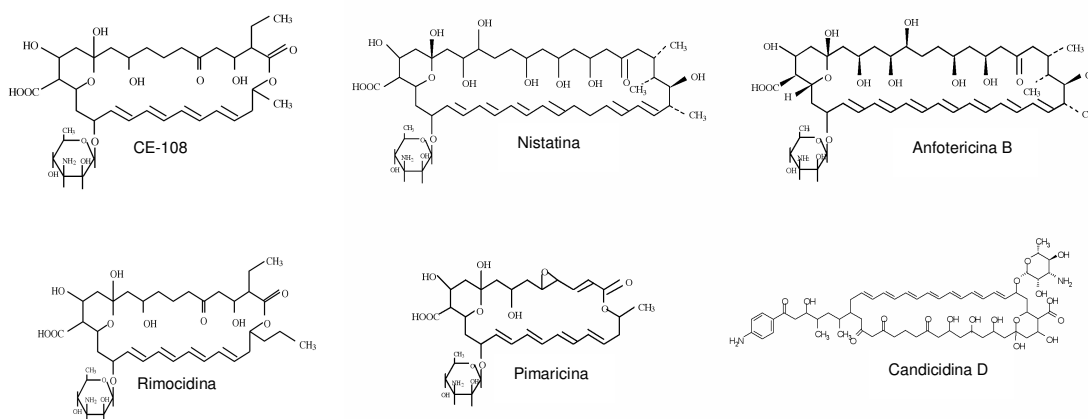


Fig. 7. Estructuras químicas de macrólidos polienos.

- **Poliquétidos poliéteres:** se trata de moléculas lineales formadas a partir de una unidad iniciadora acetato por condensación de unidades metilmalonato o etilmalonato. Dentro de este grupo destacan monensina A y B producidas por *S. cinnamomensis* (Figura 8). Igual que los macrólidos, los poliéteres son sintetizados por PKSs Tipo I.

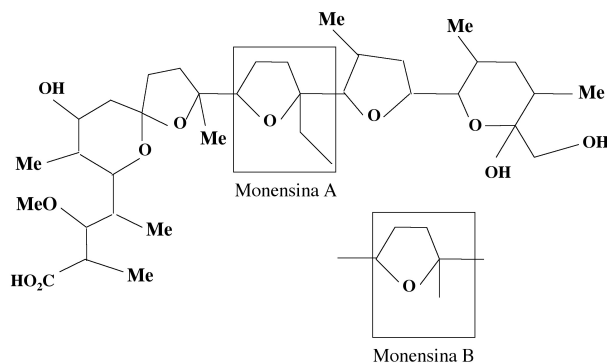


Fig. 8. Estructuras químicas de monensina A y B.

1.4. PKSs Tipo I

Organización génica

El conjunto de actividades asociadas a una PKS Tipo I se encuentra distribuido a lo largo de un número de polipéptidos multifuncionales variable que igualmente reciben el nombre de PKSs y que a su vez están compuestos por diferentes dominios catalíticos.

Los dominios catalíticos KS, AT y ACP necesarios para la biosíntesis de la cadena policetónica forman una asociación invariable, mientras que los dominios catalíticos KR, DH y ER pueden estar presentes dependiendo del grado de reducción de dicha cadena. Así se originarán grupos hidroxilo cuando el dominio de reducción presente sea KR; grupos enoilo cuando los dominios presentes sean KR y DH; y grupos C-C reducidos cuando estén presentes los tres dominios funcionales. Si no existe ninguno de los anteriores dominios catalíticos de reducción lo que aparece en la cadena policetónica son grupos ceto. El orden de los dominios catalíticos dentro de cada módulo es: KS, AT, (DH, ER, KR cuando existan) y ACP, exactamente igual que para las FASs de animales (Wakil, 1989).

A diferencia del modo de acción de las PKSs Tipo II, donde cada proteína actúa de manera reiterada en diferentes puntos de elongación de la cadena policetónica, en las PKSs Tipo I los diferentes polipéptidos que la componen actúan en un solo paso de la ruta de biosíntesis, de tal manera que existen tantos dominios catalíticos KS, AT y ACP como módulos de elongación son necesarios para constituir la cadena policetónica.

El polipéptido encargado del inicio de la biosíntesis está formado, al menos, por el módulo de "carga" y es el encargado de tomar la unidad iniciadora de la ruta biosintética. Este módulo de "carga" presenta naturalezas muy diversas. Por ejemplo en el caso de la nistatina está constituido por los dominios: KS, AT, DH y ACP; en el caso de la eritromicina está constituido por los dominios AT y ACP; en el caso de la rapamicina están presentes un dominio

ácido carboxílico-CoA ligasa y un dominio AT. Este módulo de “carga”, en el caso de los macrólidos polienos descritos, puede encontrarse tanto aislado del resto de polipéptidos multifuncionales, así ocurre para los polienos nistatina (Brautaset *et al.*, 2000), anfotericina (Caffrey *et al.*, 2001), pimaricina (Aparicio *et al.*, 1999 y 2000), rimocidina y CE-108 (Seco *et al.*, 2004) como asociado a módulos de elongación, es el caso de candicidina (Campelo y Gil, 2002) y rapamicina (Schwecke *et al.*, 1995).

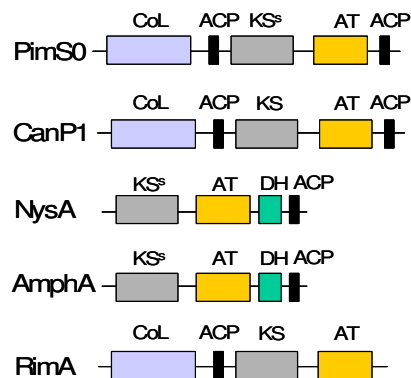


Fig. 9. Módulos de “carga” de PKSs Tipo I de los macrólidos polienos pimaricina (PimS0); candicidina (CanP1); nistatina (NysA); anfotericina B (AmphA); rimocidina y CE-108 (RimA).

El resto de PKSs encargadas de la elongación de la cadena policetónica presentan tamaños muy variables según el número de módulos de elongación que posean.

La PKS implicada en el último módulo de elongación suele presentar en su extremo C-terminal un dominio tioesterasa (TE). Este dominio es el encargado de hidrolizar el enlace entre la cadena policetónica ya formada y el brazo de 4'-fosfopantoteína del último dominio ACP. Además de esta actividad enzimática tioesterasa integrante de una PKS, existen otras proteínas tioesterasas independientes cuya función es corregir aquellos errores que se pueden producir durante el ensamblaje de la cadena (Kotowska *et al.*, 2002).

Entre las secuencias de nucleótidos que codifican para estas PKSs, existe un alto grado de conservación que ha permitido identificar una serie de motivos altamente conservados para cada uno de los dominios catalíticos que constituyen la PKS.

La secuencia conservada para los dominios KS es GPXXXXXTACSS en torno al residuo cisteína del centro activo (Bevitt *et al.*, 1992; Donadio y Katz, 1992; Aparicio *et al.*, 1996). Existen además, dos residuos histidina conservados a 135 y 173 nucleótidos del centro activo del dominio KS, en dirección al extremo C-terminal; dichos residuos histidina están siempre presentes, tanto en PKSs Tipo I como en PKSs Tipo II (Aparicio *et al.*, 1996). Respecto al dominio KS del módulo de “carga” de las PKSs de macrólidos polienos, normalmente éste presenta una sustitución del aminoácido cisteína (KS^c) del centro activo por el aminoácido serina (KS^s) (Aparicio *et al.*, 1999; Brautaset *et al.*, 2000; Caffrey *et al.*, 2001). Al principio, se pensó que dicha sustitución hacía que el dominio KS^s fuese inactivo (Aparicio *et al.*, 1999) pero más tarde se le asoció una posible actividad decarboxilasa (Brautaset *et al.*, 2000; Aparicio *et al.*, 2000), ya descrita para aquellos dominios KS en los que el aminoácido cisteína estaba sustituido por el aminoácido glutamina en el centro activo (KS^q) (Witkowski, 1999). Estudios

posteriores en la PKS que contiene el módulo de “carga” de nistatina (NysA) han confirmado esta actividad decarboxilasa para el dominio KS^s (Brautaset *et al.*, 2003).

Los dominios funcionales AT presentan la secuencia conservada GH**S**XG alrededor del residuo serina del centro activo. Igualmente existen residuos conservados glutamina, arginina e histidina en las posiciones 11, 117 y 190 respectivamente (Serre *et al.*, 1995; Aparicio *et al.*, 1996). Dentro de estos dominios catalíticos AT es posible discernir la incorporación de una unidad elongadora u otra en la cadena policetónica por la secuencia conservada de aminoácidos que en ellos aparece. Así la secuencia consenso RVDVV-----M-S-A—W-11aa-GH**S**XG determina la incorporación de metilmalonil-CoA (propionato) como unidad elongadora (Haydock *et al.*, 1995), mientras que la secuencia consenso ETGYA-----Q-A-FGLL-11aa-GH**S**XG determina que se incorpore malonil-CoA (acetato) en la elongación de la cadena policetónica en formación (Marsden *et al.*, 1994).

La secuencia consenso característica de los dominios catalíticos ACP es LGXD**S**LXXVE alrededor del residuo serina del centro activo sobre el cual se une al brazo de 4'-fosfopantoteína.

Los dominios catalíticos KR y ER presentan el motivo conservado GXGXXGXXXA en todos los casos descritos, excepto para rapamicina donde el motivo conservado es GXGXXAXXA (Scrutton *et al.*, 1990; Aparicio *et al.*, 1996).

Los dominios catalíticos DH presentan la secuencia consenso HXXXGXXXXP (Bevitt *et al.*, 1992).

Como se dijo anteriormente, los dominios funcionales están colocados dentro de cada PKS en el orden preciso en el que van a ser utilizados para la elongación de la cadena policetónica. De tal modo que si están presentes todos ellos el orden en que se van a encontrar será: KS-AT-ACP-KR-DH-ER.

Aparte de las actividades enzimáticas englobadas dentro de las PKSs y que son responsables de la elongación y posterior ciclación de la cadena policetónica, existen una serie de enzimas que una vez sintetizada la estructura básica del poliquétido se encargan de introducir una serie de modificaciones de “adorno” o post-PKS contribuyendo así, a ampliar la variabilidad de estructuras químicas que pueden ser generadas. Dentro de estas modificaciones destacan las producidas por citocromo P450 monooxigenasas, metiltransferasas, glicosiltransferasas...

Las citocromo P450 monooxigenasas son enzimas oxidativas descritas en muchos *clusters* de biosíntesis de poliquétidos macrólidos, como eritromicina (Haydock *et al.*, 1995;

Weber *et al.*, 1991; Stassi *et al.*, 1993), oleandomicina (Rodríguez *et al.*, 1995), rapamicina (Mólnar *et al.*, 1996), pimaricina (Aparicio *et al.*, 2000), nistatina (Brautaset *et al.*, 2000), rimocidina y CE-108 (Seco *et al.*, 2004) entre otros. En algunos casos, estas enzimas van asociadas a ferredoxinas, encargadas de mediar el transporte de electrones requerido por las citocromo P450 (O'Keefe y Harder, 1991).

Otro grupo de enzimas son las transferasas, capaces de reconocer una gran variedad de estructuras y transferirles diferentes grupos funcionales (metilo, glicosilo,...); en función del

grupo funcional que transfieren se pueden agrupar en: glicosiltransferasas, metiltransferasas, acetiltransferasas, aminotrasferasas, amidotrasferasas,...

Modo de acción de las PKSs Tipo I

La biosíntesis del poliquétido comienza con la PKS que constituye el módulo de “carga” encargado de tomar la unidad iniciadora y transferirla al dominio KS del primer módulo de elongación.

Las unidades iniciadoras pueden ser muy variadas. Las más frecuentes son ésteres de ácidos carboxílicos lineales (acetil-CoA, propionil-CoA o butiril-CoA) aunque están descritas unidades iniciadoras más complejas: ácido carboxílico ciclohexano, para la biosíntesis de los macrólidos no polienos FK506 y rapamicina (Motamedi y Shafiee, 1998; Aparicio *et al.*, 1996); anillos aromáticos como el ácido p-amino benzoico (PABA), para la biosíntesis del macrólido polieno candicidina (Gil *et al.*, 1985).

Es el dominio KS del primer módulo de elongación el responsable de la condensación entre la unidad iniciadora y la unidad de elongación incorporada por este módulo.

Las unidades de elongación suelen ser acetatos, propionatos o butiratos provenientes de la decarboxilación de sus correspondientes ácidos dicarboxílicos: malonato, metilmalonato y etilmalonato respectivamente. Si la unidad elongadora incorporada es metilmalonato o etilmalonato, se origina una cadena lateral metilo o etilo unida al anillo macrocíclico, debido a que la condensación tiene lugar en el carbono adyacente al grupo carboxilo.

Tras esta condensación entre la unidad iniciadora y la unidad elongadora, el residuo policetónico originado es transferido por el dominio catalítico AT al dominio catalítico ACP. Este residuo policetónico previamente puede haber sufrido algún tipo de reducción mediante la presencia de los dominios funcionales KR, DH y ER.

A medida que se van incorporando unidades de elongación mediante la acción de los diferentes polipéptidos funcionales que constituyen la PKS Tipo I, se va formando la cadena policetónica que finalmente será liberada del complejo multienzimático por acción del dominio catalítico TE. Dicho dominio se encarga de hidrolizar el enlace entre la cadena policetónica ya formada y el grupo prostético 4'-fosfopantoteína del dominio ACP del último módulo de elongación. A continuación, la cadena policetónica se cicla y es susceptible de sufrir modificaciones de “adorno” o post-PKS mediante la acción de las citocromo-P450 y de las transferasas que dan lugar a la estructura final de cada compuesto.

La primera PKS Tipo I caracterizada fue la 6-desoxieritronólido B sintasa (DEBS), implicada en la biosíntesis de la aglicona de eritromicina, 6-desoxieritronólido B (Cortés *et al.*, 1990; Donadio y Katz, 1992; Bevitt *et al.*, 1992). Esta PKS Tipo I se encuentra codificada por tres genes, *eryAI*, *eryAII* y *eryAIII* que a su vez codifican tres proteínas multifuncionales, DEBS1, DEBS2 y DEBS3.

La proteína DEBS1 está constituida por el módulo de “carga” unido a los dos primeros módulos de elongación de la cadena policetónica; la proteína DEBS2 la componen los módulos de elongación 3 y 4 y la proteína DEBS3 está formada por los módulos de elongación 5 y 6, éste último lleva asociado el dominio catalítico TE que libera la cadena policetónica formada del

complejo enzimático. Esta cadena posteriormente se cicla y por glicosilación se origina la estructura característica de eritromicina A (Figura 10).

Posteriormente, se han ido caracterizando las rutas biosintéticas de macrólidos como rapamicina, nistatina, pimarcina, candicidina, anfotericina, rimocidina, CE-108 y otros.

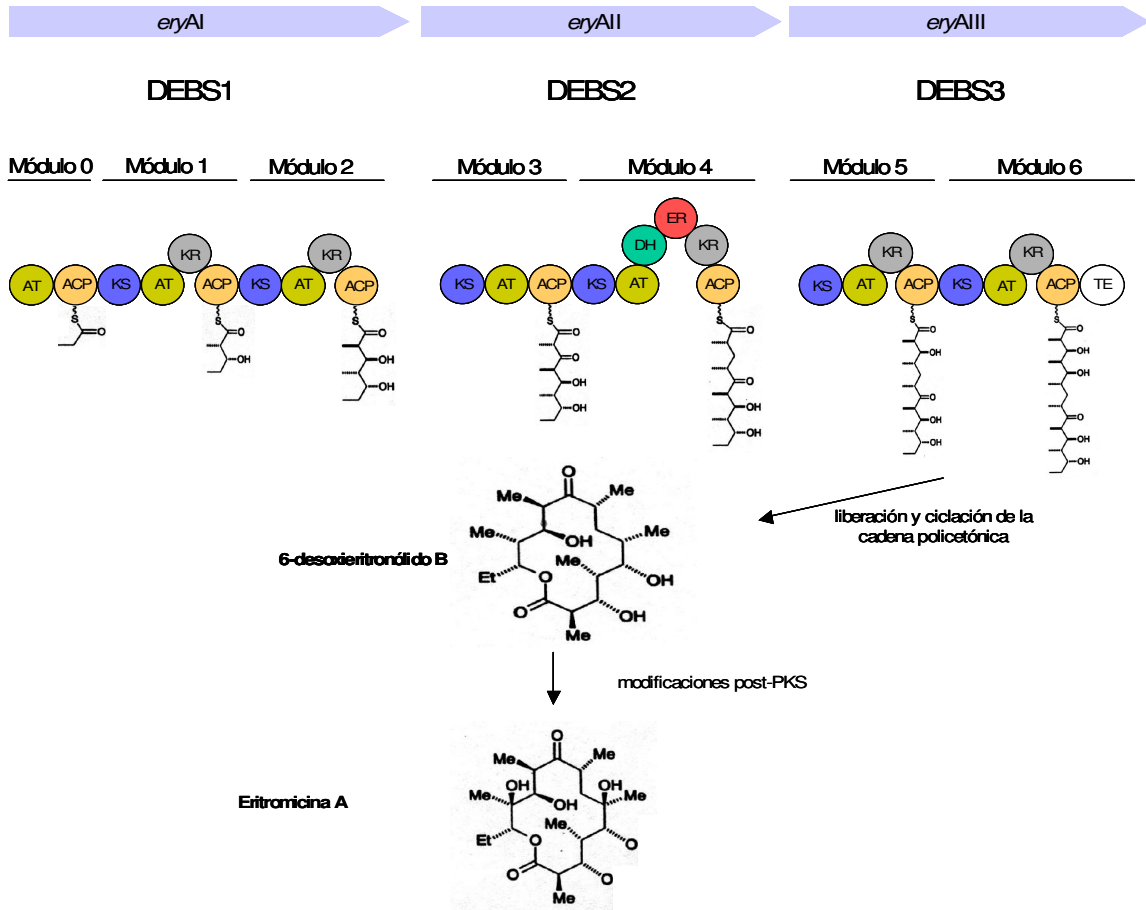


Fig. 10. Modelo de biosíntesis de eritromicina (Hopwood, 1997).

1.5. MANIPULACIÓN GENÉTICA DE RUTAS BIOSINTÉTICAS DE POLIQUÉTIDOS

Con el objetivo de descubrir y diseñar nuevas drogas, se han empleado diferentes estrategias para manipular genéticamente las rutas de biosíntesis de poliquétidos.

Debido a la estructura modular que presentan las PKSs Tipo I implicadas en la biosíntesis de antibióticos macrólidos y poliéteres, donde cada actividad enzimática está implicada en un único paso bioquímico de la ruta de biosíntesis, se puede determinar la correlación exacta entre la estructura química del macrólido o poliéter y la organización de los diferentes dominios funcionales de las distintas PKSs que constituyen dicha ruta. De tal manera que, por pérdida de una única actividad enzimática o por sustitución de determinados dominios funcionales por otros heterólogos, se pueden generar nuevos compuestos con actividad biológica.

Manipulaciones genéticas de este tipo han sido realizadas en la ruta de biosíntesis de eritromicina, donde se han obtenido nuevos compuestos bien mediante inactivación de determinados dominios funcionales en la sintasa de eritromicina (DEBS) (Donadio y Hutchinson, 1991; Donadio *et al.*, 1993); mediante reducción en el número de módulos de condensación (Cortés *et al.*, 1995; Kao *et al.*, 1995); o mediante sustitución por dominios funcionales de otras PKSs Tipo I. Este último caso, se trata de la sustitución de un dominio funcional AT presente en la sintasa DEBS1 por un dominio funcional AT de rapamicina que provoca un cambio en la especificidad por la unidad elongadora incorporada (Oliynyk, 1996; Ruan *et al.*, 1997; Stassi *et al.*, 1998).

Otro tipo de estrategia a seguir es la conocida como biosíntesis combinatorial, se trata de la expresión de los genes responsables de la PKS mínima en hospedadores heterólogos que desemboca en la formación de nuevos compuestos. Es el caso por ejemplo de la expresión del complejo DEBS1-TE de la ruta biosintética de eritromicina en el hospedador heterólogo *Streptomyces coelicolor*; dicha expresión permitió la formación de un nuevo compuesto por incorporación de acetil-CoA como unidad iniciadora en lugar de propionil-CoA como ocurre en el productor natural de eritromicina *Saccharopolyspora erythraea* (Kao *et al.*, 1995; Brown *et al.*, 1995).

1.6. ANTECEDENTES Y OBJETIVOS DEL PRESENTE TRABAJO

En la búsqueda de cepas de *Streptomyces* productoras de compuestos biocidas, se seleccionó una a partir de aislados de suelos a la que se le dio el nombre de *Streptomyces* sp. CO-AC25. Dicha cepa se seleccionó por su capacidad de producir compuestos con elevada actividad biológica frente a distintas estirpes de hongos y de bacterias. Particularmente se caracterizaba por su actividad antifúngica. Dado que otros proyectos llevados a cabo en el laboratorio estaban orientados a la caracterización y obtención de compuestos fungicidas, se consideró interesante iniciar su estudio, por lo que los objetivos propuestos fueron:

1. Caracterizar los metabolitos activos producidos por *Streptomyces* sp. CO-AC25.
2. Aislar y caracterizar los genes implicados en la biosíntesis del compuesto fungicida producido por *Streptomyces* sp. CO-AC25.
3. Manipular la ruta biosintética con el fin de generar nuevas formas moleculares no naturales que pudieran tener propiedades biológicas mejoradas respecto a la parental.

MATERIALES Y MÉTODOS

2.1. MICROORGANISMOS UTILIZADOS

Los microorganismos utilizados para el desarrollo de este trabajo se encuentran reflejados en la siguiente tabla:

Organismo	Genotipo	Referencia
<i>Escherichia coli</i> JM101	<i>SupE thiΔ(lac-proAB)</i> (F' <i>traD36 proAB lac^d</i> ZΔM15)	Yanish-Perron <i>et al.</i> , (1985)
<i>Escherichia coli</i> JM109	<i>recA1 supE44 endA1</i> <i>hsdR17 gyrA96 relA1</i> <i>thiΔ(lac-proAB)</i>	Yanish-Perron <i>et al.</i> , (1985)
<i>Escherichia coli</i> ET12567	<i>dam⁻ dcm⁻</i>	MacNeil <i>et al.</i> , (1992)
<i>Escherichia coli</i> ET12567 (pUZ8002)	<i>km^r cm^r</i>	Paget <i>et al.</i> , (1999)
<i>Escherichia coli</i> XL1-blue MRA	$\Delta(mcrA)183 \Delta(mcrCB-$ <i>hsdSMR-mrr)173</i> <i>endA1 supE44 thi1</i> <i>recA gyrA96 relA1</i> <i>lac</i>	Stratagene
<i>Escherichia coli</i> S17-1	cepa conjugativa resistente a espectinomomicina	Simon <i>et al.</i> , (1983)
<i>Streptomyces</i> sp. CO-AC25		Este trabajo
<i>Streptomyces lividans</i> TK21	SLP2 ⁻ , SLP3 ⁻	Hopwood <i>et al.</i> , (1983)
<i>Penicillium chrysogenum</i>	silvestre	ATCC 8537
<i>Micrococcus luteus</i>	silvestre	ATCC 4698

Tabla 4. Microorganismos utilizados en este trabajo.

2.2. CRECIMIENTO Y CONSERVACIÓN

Las cepas de *Escherichia coli* y *Micrococcus luteus* se crecieron a 37°C tanto en medio sólido como líquido; en este caso, a una agitación de 250 rpm y se conservaron como suspensiones celulares en glicerol 20% a una temperatura de -20°C. El crecimiento de *E. coli* se determinó midiendo la turbidez de los cultivos (absorbancia a 600 nm) en el espectrofotómetro. Las cepas de *E. coli* utilizadas más frecuentemente también se conservaron mediante resiembra periódica en placas de agar con el medio de cultivo adecuado y los suplementos necesarios. Las placas se mantuvieron selladas con Parafilm® a 4°C.

Las cepas de *Streptomyces* se crecieron en los diferentes medios de cultivo a la temperatura de 30°C y a una agitación de 250 rpm para los medios líquidos. En algunos casos la agitación se realizó con bolas de vidrio y muelles para facilitar tanto la oxigenación como la dispersión del micelio. Se conservaron como suspensiones de esporas en glicerol 20% a una temperatura de -20°C. Para obtener las esporas de *S. lividans* TK21 se sembró el microorganismo en placas de R2YE adecuadamente suplementado (Materiales y Métodos, apartado 2.3); las esporas de *S. sp.* CO-AC25 se obtuvieron tanto a partir de placas de R2YE adecuadamente suplementado (Materiales y Métodos, apartado 2.3) como a partir de placas de MS suplementadas con MgCl₂ 10 mM (Materiales y Métodos, apartado 2.3). En todos los casos, las placas se incubaron a 30°C durante un tiempo estimado de 6-7 días. Las esporas se recogieron con la ayuda de un asa de siembra estéril, se filtraron a través de un filtro de algodón hidrófilo estéril, se centrifugaron 10 minutos a 4000 rpm, se lavaron con H₂O estéril y se resuspendieron en un volumen adecuado de glicerol 20% para su posterior congelación (Hopwood *et al.*, 1985). Las cepas de *Streptomyces* también se conservaron como micelio congelado, obtenido a partir de cultivo líquido.

La cepa *Penicillium chrysogenum* se creció sobre medio PDA a 30°C hasta su esporulación y se conservó como suspensión de esporas en glicerol 20% a una temperatura de -20°C.

2.3. MEDIOS DE CULTIVO

Los medios de cultivo utilizados para el crecimiento de *Streptomyces* fueron:

YEME: Medio de cultivo líquido usado para la extracción de plásmidos y ADN total de *Streptomyces lividans* TK21. (Hopwood *et al.*, 1985).

YES: Medio de cultivo líquido para la obtención de ADN total de *Streptomyces sp.* CO-AC25. Su composición es: extracto de levadura 1%, sacarosa 10,3%. Tras ajustar el pH a 7,2 con NaOH y autoclavar se suplementa con glucosa 1%. (R. Santamaría, comunicación personal).

TSB (Tryptone Soya Broth 3%): Medio de cultivo utilizado tanto en la conjugación entre *Streptomyces sp.* CO-AC25 y *E. coli* S17-1, como en la electroporación de *S. sp.* CO-AC25. (Hopwood *et al.*, 1985).

E1: Medio de cultivo usado en la fermentación para la obtención de las actividades biocidas producidas por *Streptomyces sp.* CO-AC25. Su composición es: Pharmamedia 0,5%, almidón 2%, glucosa 2%, extracto de levadura 0,25%, NaCl 0,3%, CaCO₃ 0,3%, MgSO₄•7H₂O 0,1%, KH₂PO₄ 0,1%. Tras ajustar el pH a 7,4 con NaOH se añade la resina XAD-16 2% y se autoclava.

SYM2: Se utilizó como medio de cultivo para la obtención de protoplastos de *Streptomyces* sp. CO-AC25. Está compuesto por: extracto de carne 0,25%, triptona 0,4%, peptona de soja 0,5%, glucosa 0,2%, almidón 2,5%, extracto de levadura 0,3%. Tras ajustar el pH a 7,0 con NaOH, se autoclava y se suplementa con glicina 0,5%.

R2YE (R5): Este medio de cultivo fue utilizado para la transformación, transfección y conjugación, así como medio de esporulación de las distintas cepas de *Streptomyces*. En el caso de *Streptomyces* sp. CO-AC25 también se usó ocasionalmente como medio de producción de actividades biocidas. (Thompson *et al.*, 1980).

MS (SFM): Se utilizó como medio selectivo tras la conjugación entre *E. coli* y *Streptomyces* sp. CO-AC25 y también como medio de esporulación de *S. sp.* CO-AC25. Su composición es: manitol 2%, harina de soja 2%, agar 2%. Se autoclava dos veces. (Hobbs *et al.*, 1989).

CRM: Medio utilizado en los ensayos de electroporación de *Streptomyces* sp. CO-AC25. (Pigac y Schrempf, 1995).

DNB (Difco Nutrient Broth 0,8%): Se utilizó para la preparación de suspensiones de fagos de *Streptomyces*. (Hopwood *et al.*, 1985).

SNA: Agar de cobertura. (Hopwood *et al.*, 1985).

DNA: Medio utilizado para la obtención de alto título y ADN de actinofagos. (Hopwood *et al.*, 1985).

Los medios de cultivo utilizados para el crecimiento de *E. coli* fueron:

LB (Luria-Bertani): Medio utilizado para el cultivo de *E. coli*. (Sambrook *et al.*, 1989). Cuando fue preciso se suplementó con indicadores de complementación del gen *lacZ*, como X-Gal 40 µg/ml e IPTG 12,5 µg/ml y los correspondientes marcadores de selección.

2xTY: Medio de cultivo utilizado para crecer *E. coli* S17-1 en la conjugación con *Streptomyces* sp. CO-AC25 (Maniatis *et al.*, 1982). Su composición es: bactotripton 2%, extracto de levadura 1%.

Medio mínimo M9: Medio de cultivo utilizado para el aislamiento y conservación de cepas de *E. coli* (Miller, 1972).

Otros medios de cultivo utilizados fueron:

Medio PDA: Potato Dextrose Agar (39 g/L). Utilizado para obtener esporas de *Penicillium chrysogenum*.

2.4. SELECCIÓN CON ANTIBIÓTICOS

Los antibióticos utilizados para la selección de cepas resistentes tanto de *Streptomyces* como de *E. coli* se muestran en la tabla 5.

Antibiótico	Concentración de selección para <i>Streptomyces</i>		Concentración de selección para <i>E. coli</i>
	Medio sólido	Medio líquido	Medios sólido y líquido
Ampicilina			100
Cloranfenicol			35
Tetraciclina			10
Apramicina	50	50	100
Espectinomicina			25
Kanamicina	50	10	50
Ácido nalidíxico			20
Thiostrepton	50	10	
Eritromicina	200	100	
Gentamicina	100	100	

Tabla 5. Concentraciones de selección en µg/ml en medio sólido y en medio líquido de los distintos antibióticos para *Streptomyces* y *E. coli*.

La preparación de estos antibióticos se realizó, dependiendo del antibiótico, bien mediante disolución en agua y posterior esterilización por filtración con filtros Millipore® 0,45µm, o mediante disolución en etanol o DMSO, en cuyo caso no precisaron de otro tipo de esterilización.

2.5. VECTORES

Los vectores usados para la realización de este trabajo se encuentran reflejados en la siguiente tabla:

Vector	Características	Referencia
pIJ922	Plásmido de <i>Streptomyces</i> de bajo número de copias, <i>tsr^r</i>	Lydiate <i>et al.</i> , (1985)
pIJ941	Plásmido de <i>Streptomyces</i> de bajo número de copias, <i>tsr^r</i> , <i>hyg^r</i>	Lydiate <i>et al.</i> , (1985)
pIJ486	Plásmido de <i>Streptomyces</i> de alto número de copias, <i>tsr^r</i> , <i>neo^r</i>	Ward <i>et al.</i> , (1986)
pGM9	Plásmido de <i>Streptomyces</i> de alto número de copias, <i>tsr^r</i> , <i>neo^r</i>	Muth <i>et al.</i> , (1989)
pCNB3070	Plásmido de <i>Streptomyces</i> de alto número de copias, <i>tsr^r</i>	Fernández- Moreno <i>et al.</i> , (1992) (comunicación personal)
KC212	Actinofago derivado de ϕ C31, <i>ap^r</i> , (<i>att^r</i>)	Bruton C.J. <i>et al.</i> , (1991)
PM1	Actinofago derivado de ϕ C31, <i>tsr^r</i> , <i>hyg^r</i> (<i>att^r</i>)	Malpartida y Hopwood, (1986)
PM1(Apra)	Actinofago derivado de ϕ C31, <i>ap^r</i> , (<i>att^r</i>)	Este trabajo
pIJ2925	Plásmido de <i>E.coli</i> para el clonaje, <i>amp^r</i> , <i>βgal</i>	Janssen y Bibb, (1993)
pUC19	Plásmido de <i>E.coli</i> para el clonaje, <i>amp^r</i> , <i>βgal</i>	Yanish-Perron <i>et al.</i> , (1985)
pUC18	Plásmido de <i>E.coli</i> para el clonaje, <i>amp^r</i> , <i>βgal</i>	Yanish-Perron <i>et al.</i> , (1985)
pSU20	Plásmido de <i>E.coli</i> para el clonaje, <i>cm^r</i> , <i>βgal</i>	Bartolome <i>et al.</i> , (1991)
pBR329	Plásmido de <i>E.coli</i> para el clonaje, <i>amp^r</i> , <i>cm^r</i> , <i>Tc^r</i>	Covarrubias y Bolivar., (1982)
EMBL-4	Vector de <i>cloning</i> de <i>E. coli</i> basado en el Fago λ	Frischauf <i>et al.</i> , (1983)
pOJ260	Plásmido bifuncional utilizado en la conjugación entre <i>E. coli</i> y <i>Streptomyces</i> , <i>ap^r</i> , <i>rep^{UC}</i>	Bierman <i>et al.</i> , (1992)
pSET152	Plásmido bifuncional utilizado en la conjugación entre <i>E. coli</i> y <i>Streptomyces</i> , <i>ap^r</i> , <i>rep^{UC}</i> , <i>int^{ϕC31}</i> , <i>βgal</i>	Bierman <i>et al.</i> , (1992)
pHJL401	Plásmido bifuncional basado en los orígenes de replicación del pUC19 y SCP2* de <i>Streptomyces</i> , <i>amp^r</i> , <i>tsr^r</i> .	Larson y Hershberger, (1986)

Tabla 6. Plásmidos y fagos usados como vectores en *Streptomyces* y *E. coli*.

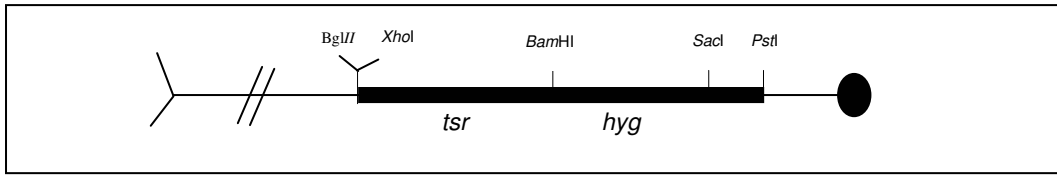


Fig . 11. Mapa físico del vector PM1.

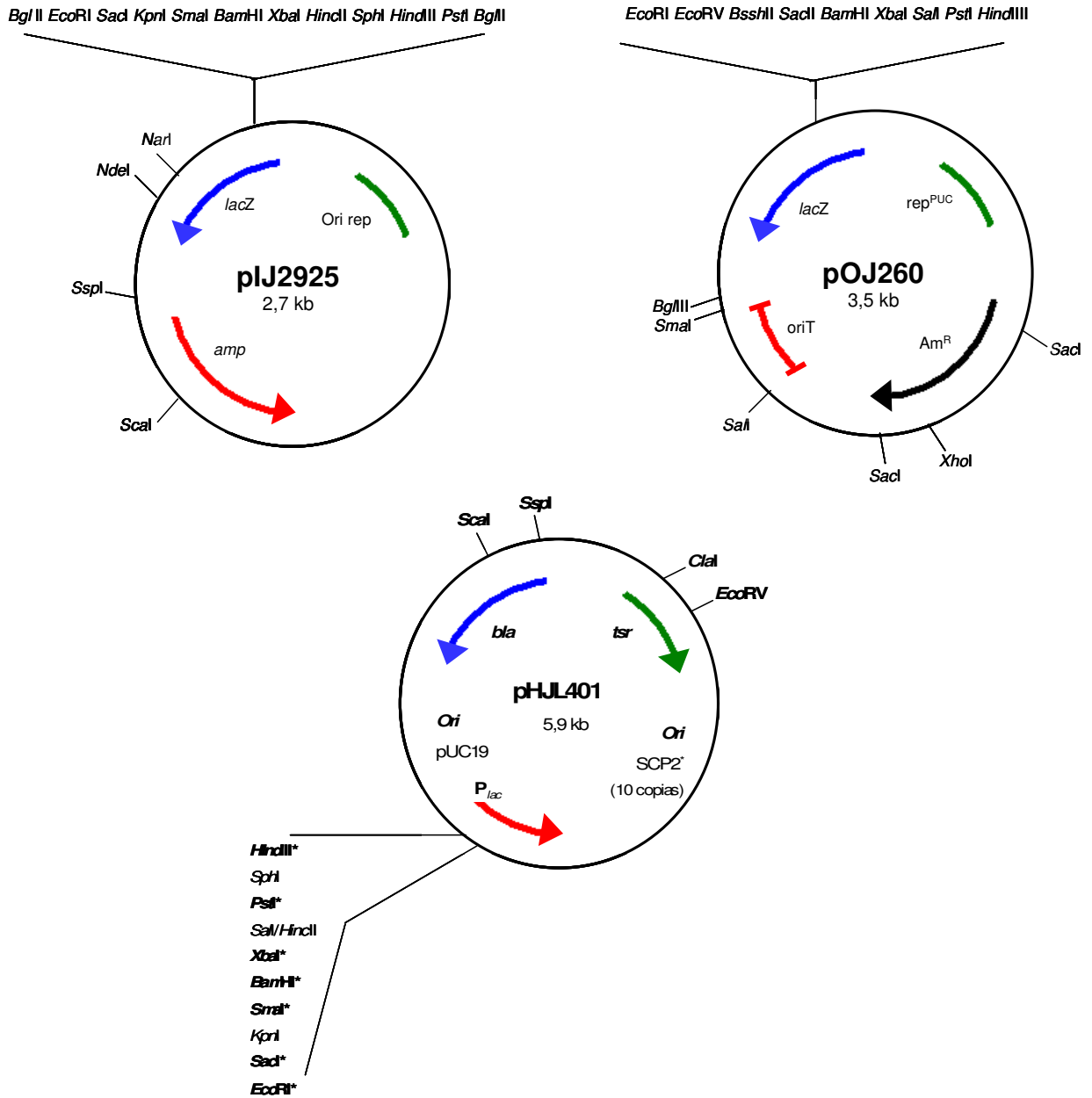


Fig. 12. Mapa físico de los vectores pIJ2925, pOJ260 y pHJL401.

Otros plásmidos contruidos para la realización de este trabajo se muestran en las siguientes tablas:

Plásmido	Descripción / Uso	Referencia
pSM23.8	Plásmido de <i>E. coli</i> derivado del vector pIJ2925 que contiene un fragmento <i>SacI</i> de 3,1 kb y que codifica el módulo de arranque para la PKS de biosíntesis de rimocidina. Usado como sonda para rastrear la librería genómica de <i>S. sp.</i> CO-AC25.	Seco <i>et al.</i> , (2004)
pAS11J2-1	Plásmido de <i>E. coli</i> derivado del vector pIJ2925, que contiene un fragmento interno <i>SacI/BamHI</i> de 3,1 kb aislado del cromosoma de <i>S. sp.</i> CO-AC25 proveniente del fago λ EMBL-4/11J. Sirvió como paso previo a la construcción del plásmido pLM31.	Este trabajo
pFPZ12	Plásmido de <i>E. coli</i> derivado del vector pIJ2925, que contiene un fragmento interno <i>SacI</i> de 2,9 kb aislado del cromosoma de <i>S. sp.</i> CO-AC25 proveniente del fago λ EMBL-4/1B. Paso previo a la construcción del pLM32.	Este trabajo
pTC8	Plásmido de <i>E. coli</i> derivado del vector pIJ2925, que contiene un fragmento interno <i>PstI</i> de 4,5 kb aislado del cromosoma de <i>S. sp.</i> CO-AC25 proveniente del fago λ EMBL-4/P1. Paso previo a la construcción del pLM33.	Este trabajo

Tabla 7. Plásmidos usados en este trabajo.

Plásmido	Descripción / Uso	Referencia
pLM31	Plásmido bifuncional derivado del vector pOJ260, que contiene el fragmento interno <i>SacI/BamHI</i> de 3,1 kb proveniente del clon pAS11J2-1, usado para realizar la disrupción génica.	Este trabajo
pLM32	Plásmido bifuncional derivado del vector pOJ260, que contiene el fragmento interno <i>SacI</i> de 2,9 kb proveniente del clon pFPZ12, usado para realizar la disrupción génica.	Este trabajo
pLM33	Plásmido bifuncional derivado del vector pOJ260, que contiene el fragmento interno <i>PstI</i> de 4,5 kb proveniente del clon pTC8, usado para realizar la disrupción génica.	Este trabajo
pAS11J7	Plásmido de <i>E. coli</i> derivado del vector pIJ2925 que contiene un fragmento <i>BamHI</i> de 2 kb aislado del cromosoma de <i>S. sp.</i> CO-AC25 utilizado para rastrear la genoteca y aislar el fago λ EMBL-4/F1.	Este trabajo
pAS13	Plásmido de <i>E. coli</i> derivado del vector pIJ2925 que contiene un fragmento <i>SacI/EcoRI</i> de 2,2 kb aislado del cromosoma de <i>S. sp.</i> CO-AC25 utilizado para rastrear la genoteca y aislar el fago λ EMBL-4/Db.	Este trabajo

Tabla 8. Plásmidos usados en este trabajo.

Plásmido	Descripción / Uso	Referencia
pG45	Plásmido de <i>E. coli</i> derivado del vector pIJ2925 que contiene un fragmento <i>SphI</i> de 2 kb aislado del cromosoma de <i>S. sp.</i> CO-AC25 utilizado para rastrear la genoteca y aislar el fago λ EMBL-4/b.	Este trabajo
pFPZ15	Plásmido de <i>E. coli</i> derivado del vector pIJ2925 que contiene un fragmento <i>BamHI-EcoRI</i> de 600 pb aislado del cromosoma de <i>S. sp.</i> CO-AC25 utilizado para rastrear la genoteca y aislar el fago λ EMBL-4/K.	Este trabajo
pK2.1	Plásmido de <i>E. coli</i> derivado del vector pIJ2925 que contiene un fragmento <i>BglI/EcoRI</i> de 400 pb aislado del cromosoma de <i>S. sp.</i> CO-AC25 utilizado para rastrear la genoteca y aislar el fago λ EMBL-4/G10.	Este trabajo

Tabla 9. Plásmidos usados en este trabajo.

2.6. AISLAMIENTO DE ADN

Aislamiento de ADN plasmídico de *E. coli*

Dicho aislamiento se realizó siguiendo el método de lisis alcalina descrito por Birnboim (1983).

Para la preparación de aquellas muestras que posteriormente fueron secuenciadas se utilizó el kit Plasmid Miniprep de Biorad (Ref.: 732-6100).

Aislamiento de ADN plasmídico de *Streptomyces*

La extracción de ADN plasmídico de *Streptomyces* también se realizó siguiendo el método de la lisis alcalina (Birnboim, 1983). Como paso previo a la lisis se incubó el micelio durante 30 minutos a 37°C en presencia de 4 mg/ml de lisozima (Hopwood *et al.*, 1985).

Aislamiento de ADN total de *Streptomyces*

La extracción de ADN total de *Streptomyces* se realizó según el método descrito por Hopwood *et al.*, (1985).

Para calcular la concentración del ADN total obtenido se midió su absorbancia a 260 nm:

$$1 \text{ unidad de } A_{260 \text{ nm}} = 50 \mu\text{g/ml de ADN.}$$

La pureza del ADN obtenido se estimó por el ratio $A_{260\text{nm}} / A_{280\text{nm}}$

Aislamiento de ADN de fagos de *E. coli*

La obtención de ADN procedente del fago λ EMBL-4 y los distintos recombinantes se realizó siguiendo el protocolo descrito por Sambrook *et al.*, (1989). La cepa de *E. coli* utilizada para el aislamiento del ADN de dichos fagos fue *E. coli* XL1-blue MRA (Stratagene).

Aislamiento de ADN de fagos de *Streptomyces*

La preparación de ADN de actinofagos se realizó siguiendo el protocolo descrito por Hopwood *et al.*, (1985). Como cepa infectiva se empleó *Streptomyces lividans* TK21.

2.7. OBTENCIÓN DE ALTO TÍTULO DE FAGOS

Obtención de alto título de fagos de *E. coli*

La obtención de alto título del fago λ EMBL-4 y de los distintos recombinantes se realizó siguiendo el protocolo descrito por Sambrook *et al.*, (1989).

Obtención de alto título de fagos de *Streptomyces*

La preparación de alto título de actinofagos se realizó siguiendo el protocolo de Hopwood *et al.*, (1985).

2.8. INTRODUCCIÓN DE ADN

Transformación de *E. coli*

Se realizó utilizando células competentes de *E. coli* JM101 y JM109 siguiendo el protocolo descrito por Brown *et al.*, (1979). Las células competentes una vez preparadas y alicuotadas se conservaron en glicerol a -80°C.

Transformación de *Streptomyces*

En algunos casos se realizó mediante el uso de protoplastos de *Streptomyces*. En el caso de *Streptomyces lividans* TK21 se prepararon suspensiones de protoplastos y se transformaron con distintos plásmidos siguiendo el protocolo descrito por Hopwood *et al.*, (1985).

En el caso de *Streptomyces* sp. CO-AC25 se llevaron a cabo modificaciones del protocolo descrito por Hopwood *et al.*, (1985). Se ensayaron diferentes condiciones para la obtención de protoplastos; dicha cepa se creció en distintos medios de cultivo líquidos con el fin de obtener un micelio idóneo para la formación de protoplastos, se probaron diferentes concentraciones tanto de glicina como de magnesio añadidos al medio de cultivo y distintos tiempos de incubación de los cultivos. También se ensayaron distintas concentraciones y tiempos de incubación con lisozima durante la digestión del micelio.

Transfección de *Streptomyces*

Se siguió el método descrito por Hopwood *et al.*, (1985).

Infección en *Streptomyces*

Las suspensiones de fagos con alto título a partir de *S. lividans* TK21, se utilizaron para infectar esporas de *S. sp.* CO-AC25 según el método descrito por Hopwood *et al.*, (1985).

Conjugación entre *E. coli* y *Streptomyces*

El protocolo seguido fue una modificación del propuesto por Mazodier *et al.*, (1989).

Se transformaron células competentes S17-1 (Brown *et al.*, 1979) con el plásmido con el cuál se quiere llevar a cabo la conjugación y se inoculó una colonia en medio líquido selectivo, en presencia de espectinomicina y apramicina (antibiótico requerido por la presencia del plásmido conjugativo). Se creció durante toda la noche, en agitación a 37°C.

En medio líquido se sembró la cepa de *Streptomyces* que deseábamos transformar, en este caso *S. sp.* CO-AC25, con antelación suficiente para que coincidiera con el cultivo de *E. coli* S17-1. Se creció durante 48 horas a 30°C, en agitación.

Pasado este tiempo, se hicieron diluciones de ambos cultivos en medio TSB: 1/100 para el cultivo de *E. coli* S17-1 y 1/10 para el de *S. sp.* CO-AC25 y se crecieron durante 3 horas a 37°C y 30°C respectivamente. Posteriormente se centrifugaron ambas diluciones y se resuspendieron en 1 ml de TSB. Tras hacer diluciones seriadas del cultivo de *Streptomyces* sp. CO-AC25 (cultivo receptor) se mezclaron con partes iguales del cultivo de *E. coli* S17-1 (cultivo

donador). Las mezclas se incubaron durante 3 horas a 30°C, posteriormente se plaquearon en medio sólido R5 y se incubaron durante 16 horas a 30°C, tiempo al cual se procedió a la selección con los antibióticos adecuados, apramicina requerida por el plásmido conjugativo y ácido nalidíxico. Las placas se incubaron a 30°C hasta la aparición de exconjugantes.

2.9. MANIPULACIÓN ENZIMÁTICA DEL ADN

Digestión del ADN

Las endonucleasas de restricción con las que se llevaron a cabo las digestiones de ADN se utilizaron siguiendo las recomendaciones de las distintas casas comerciales.

Desfosforilación de extremos cohesivos

Este procedimiento se realizó cuando fue preciso evitar la recircularización de aquellos vectores digeridos con endonucleasas de restricción y que iban a ser utilizados para posteriores ligaciones. Para la desfosforilación, se trataron dichos vectores con fosfatasa alcalina, enzima que hidroliza los grupos fosfato del extremo 5' del ADN impidiendo la formación de los enlaces fosfodiéster en la misma molécula, y así favorecer la incorporación de fragmentos de ADN que contengan los extremos 5' intactos.

La reacción de desfosforilación se llevó a cabo añadiendo una unidad enzimática de fosfatasa alcalina de intestino de ternera (CIAP) (Boehringer Mannheim; Ref.: 1097 075) por cada 50 picomoles de extremos 5' de ADN e incubando a 37°C durante una hora (Maniatis *et al.*, 1982). Finalmente, la fosfatasa alcalina se inactivó por calor a 65°C durante 10 minutos.

Relleno de extremos protuberantes de ADN

Se siguió el protocolo descrito por Sambrook *et al.*, (1989) utilizando el fragmento Klenow de la ADN polimerasa I de *E. coli* para rellenar los extremos de ADN 5' protuberantes y T4 ADN polimerasa para romper fragmentos con extremos 3' protuberantes. En caso de estar presentes extremos 5' y 3' protuberantes en una misma reacción se utilizó la enzima T4 ADN polimerasa, que también presenta actividad exonucleasa a partir del extremo 5'. Todas las reacciones se llevaron a cabo en las condiciones recomendadas por las casas comerciales y cuidando que el tiempo de incubación no sobrepasara los 30 minutos a 37°C para evitar posibles actividades exonucleasas no deseadas. Ambas enzimas se inactivaron a 65°C durante 10 minutos.

Ligación de fragmentos de ADN

Se utilizó la ADN ligasa del bacteriofago T4 y se tuvieron en cuenta las recomendaciones de Sambrook *et al.*, (1989), poniendo siempre un exceso molar del fragmento de ADN a clonar del orden de cuatro veces respecto al vector. Las reacciones fueron llevadas a cabo a temperatura ambiente durante 2 horas para extremos protuberantes. En el caso de ligación de extremos romos, la reacción fue llevada a cabo a 16°C durante toda la noche.

Amplificación de ADN por PCR

Todas las amplificaciones de ADN por reacción de la polimerasa en cadena (PCR) se llevaron a cabo con la siguiente mezcla de reacción: 100 ng de ADN molde; 10 picomoles de cada oligonucleótido; 0,2 mM de cada dideoxidonucleótido; 1,5 mM MgCl₂; 20 mM Tris-HCl pH8.

El programa de amplificación utilizado fue el que se describe:

1 ciclo de 5 minutos a 95°C.

5 ciclos de 2 minutos a 97°C; 1 minuto a la T_m y 1-2 minutos a 72°C.

20 ciclos de 2 minutos a 95°C; 1 minuto a la T_m; 1-2 minutos a 72°C.

1 ciclo de 7 minutos a 72°C.

La T_m osciló entre 50°C y 65°C en función de la temperatura de fusión propia de los oligonucleótidos. Los moldes utilizados fueron tanto plásmidos como ADN total de *Streptomyces*. El tiempo de extensión a 72°C osciló entre 1 y 2 minutos en función del tamaño del ADN a amplificar.

Los oligonucleótidos empleados en este trabajo fueron suministrados por la empresa Isogen. Algunos de estos oligonucleótidos se muestran en la siguiente tabla:

Oligonucleótidos	Secuencia (5'-3')
flvI	TGT GTC CTC GTT CGG CGT CA
flvII	CGA CCA CAC CCC GGT CAC AT
flvIII	TCG CGC TGT TCG ACG CCG CGA
flvIV	AAG CCG GCC TCC GCG TCG T

Tabla 10. Ejemplo de algunos oligonucleótidos diseñados para este trabajo.

2.10. ELECTROFORESIS DE ADN

Electroforesis en geles de agarosa

Las moléculas de ADN con una cantidad de tampón de carga apropiada, se separaron mediante electroforesis en geles de agarosa preparados en tampón TBE, a los que se añadió bromuro de etidio a una concentración final de 0,5 µg/ml. La concentración de agarosa utilizada para la preparación de los geles varió entre 0,5% y 2% (p/v) dependiendo del tamaño de los fragmentos de ADN a separar (Sambrook *et al.*, 1989).

Tampón TBE: Tris-HCl 89 mM, ácido bórico 89 mM, EDTA 2 mM pH8.

Tampón de carga para ADN (5X): azul bromofenol 0,25%, sacarosa 40%, Tris-HCl 10 mM pH8, EDTA 10 mM.

Bromuro de Etidio: solución 10 mg/ml en agua. Se conserva a 4°C, protegido de la luz.

Purificación de fragmentos de ADN a partir de geles de agarosa

Tras cortar del gel de agarosa el fragmento de ADN que se quiere purificar, se electroeluyó en tampón TBE 1X en bolsas de diálisis (Maniatis *et al.*, 1982). A continuación, se trató con fenol:cloroformo neutro y se precipitó en presencia de acetato sódico 0,3 M y un volumen final de isopropanol.

En determinadas ocasiones la purificación de fragmentos de ADN desde geles de agarosa se llevó a cabo utilizando el kit Perfectprep Gel Cleanup de Eppendorf (Ref.: 0032007.740).

2.11. HIBRIDACIÓN DE ADN

Transferencia de ADN desde geles de agarosa: *Southern Blot*

La transferencia de ADN fue realizada según la técnica descrita por Southern (Southern, 1975). El ADN fue transferido desde geles de agarosa a membranas de nylon (Hybond-N, Amersham Pharmacia). En el caso de fragmentos de ADN a transferir muy grandes, como es el caso del ADN total de *Streptomyces* o el ADN de fagos λ EMBL-4, el gel se irradió con luz ultravioleta de onda corta durante 2-3 minutos con objeto de fragmentar el ADN y así facilitar su transferencia.

Una vez realizada la transferencia, la fijación del ADN a la membrana de nylon se llevó a cabo tanto por exposición de la misma a luz ultravioleta durante 3 minutos como por vacío a 80°C durante 2 horas.

Hibridación de ADN

La membrana de nylon con el ADN transferido y fijado fue prehibridada, hibridada y posteriormente lavada siguiendo el protocolo descrito por Sambrook *et al.*, (1989). Las soluciones empleadas en cada caso fueron:

Solución de prehibridación: SSC5X, Denhardt 5X, SDS 0,1%, 200 μ g/ml de ADN desnaturalizado de esperma de salmón.

Solución Denhardt 50X: Albúmina de suero bovino (BSA fracción V) 1%, Ficoll (tipo 400) 1% y polivinilpirrolidona 1%. Una vez disuelto se pasó por el filtro Millipore® 0,45 μ m y se almacenó a -20°C.

Soluciones de lavado I, II y III: SSC 2X, SDS 0,1%; SSC 0,8%, SDS 0,1%; SSC 0,2%, SDS 0,1%, respectivamente.

2.12. MARCAJE RADIOACTIVO DE FRAGMENTOS DE ADN

Marcaje por desplazamiento de mellas (*Nick Translation*)

Se utilizó el método descrito por Sambrook *et al.*, (1989) basado en la acción combinada de la ADNasa I y la ADN polimerasa I de *E. coli*. La ADNasa I añadida en una concentración adecuada provoca la rotura al azar en distintos puntos del ADN que después son desplazados por la ADN polimerasa I en base a su actividad exonucleasa 5´-3´ y rellenados gracias a su actividad polimerasa 5´-3´ incorporando el nucleótido radiactivo. Como isótopos radiactivos se han utilizado [$\alpha^{32}\text{P}$]dCTP y [$\alpha^{32}\text{P}$]dGTP debido al mayor contenido en *Streptomyces* de G+C.

En ocasiones se utilizó el kit Rediprime II Random Prime Labelling System de Amersham Biosciences (Ref.: RPN1633).

Marcaje mediante el relleno de extremos protuberantes

Método igualmente descrito por Sambrook *et al.*, (1989), donde se ha aprovechado la actividad del fragmento Klenow de la ADN polimerasa I de *E. coli* para rellenar extremos 5´ protuberantes incorporando isótopos radiactivos.

En ambos casos, una vez que se obtuvo la sonda marcada radiactivamente fue purificada a través de una columna de Sephadex con el fin de eliminar aquellos nucleótidos que no se hubieran incorporado. Las columnas empleadas a tal efecto fueron columnas Micro Bio Spin P30 Tris Chromatography Columns de Biorad (Ref.: 732-6223).

2.13. SECUENCIACIÓN DE ADN

La secuenciación de las regiones del cromosoma de *Streptomyces* sp. CO-AC25 presentadas en este trabajo se llevó a cabo mediante clonación de fragmentos de ADN en plásmidos. En el Servicio de Secuenciación del Centro de Investigaciones Biológicas (CIB) y del Centro Nacional de Biotecnología (CNB) es donde se secuenciaron mediante el método de terminación de cadena por didesoxinucleótidos (Sanger *et al.*, 1977). Los oligonucleótidos usados para dicho fin fueron el M13 universal (-47) y el M13 reverso (-48).

2.14. AISLAMIENTO Y MANIPULACIÓN DE ARN

Aislamiento de ARN total de *Streptomyces*

La extracción de ARN total de *Streptomyces* se realizó introduciendo alguna modificación al método descrito por Hopwood *et al.*, (1985).

Las diferentes cepas de *Streptomyces* de las cuales se quería aislar el ARN se sembraron en placas de R5 sólido sobre celofán, con el fin de facilitar la recogida del micelio.

La extracción del ARN se llevó a cabo utilizando una mezcla de Kirby y fenol:cloroformo 6% (v/v) y tras una primera precipitación se realizó una precipitación selectiva en presencia de acetato sódico 4 M, pH6 a -20°C durante toda la noche. Una vez extraído el ARN se almacenó disuelto en agua en pequeñas alícuotas a -80°C .

El ARN total obtenido se cuantificó midiendo la absorbancia a 260 nm:

$$1 \text{ unidad de } A_{260 \text{ nm}} = 40 \mu\text{g/ml de ARN.}$$

La pureza del ARN obtenido se estimó por el ratio $A_{260 \text{ nm}} / A_{280 \text{ nm}}$.

Electroforesis de ARN en geles desnaturalizantes

El análisis del ARN aislado se llevó a cabo siguiendo el método descrito por Sambrook *et al.*, (1989).

Dicho análisis se realizó en geles de agarosa 1% en tampón MOPS y formaldehído 6,3% como agente desnaturalizante, cuyo objetivo es evitar la formación de estructuras secundarias en el ARN. La electroforesis se realizó igualmente en tampón MOPS y las muestras se prepararon tal y como se describe a continuación: x μg de ARN; 2 μl de MOPS 10X; 3,5 μl de formaldehído 37%; 10 μl de formamida 100% y $\text{H}_2\text{O-DEPC}$ hasta 20 μl .

Las muestras se calentaron durante 10 minutos a 65°C y rápidamente se enfriaron en hielo durante 2-3 minutos. Tras este tiempo se les añadió 2,2 μl de tampón de carga para ARN y se cargaron en el gel.

Tampón MOPS 10X: MOPS 0,2 M, AcNa 0,05 M, EDTA 10 mM, pH7.

2.15. ENSAYOS DE PROTECCIÓN A ENDONUCLEASA S1

Ensayos de protección a endonucleasa S1 de alta resolución

Esta técnica se realizó siguiendo el protocolo descrito por Sambrook y Russell (2001), y permite identificar el origen de transcripción de un gen.

50 μg de ARN se hibridaron con 50000 cpm de sonda marcada con $[\gamma^{32}\text{P}]\text{dATP}$ durante toda la noche a la temperatura de 62°C - 64°C (determinada empíricamente), tras incubar las muestras a 85°C durante 10 minutos. Pasado este tiempo se procedió al tratamiento de las muestras con 50U de endonucleasa S1, durante 1 hora a 37°C y en presencia de 300 μl de tampón S1. Los fragmentos protegidos se analizaron en geles de acrilamida-urea 6% junto a la reacción de secuencia de Maxam y Gilbert (1980) de la sonda.

Solución de hibridación: formamida 80%, PIPES 40 mM pH6,4, NaCl 400 mM, EDTA 1 mM pH8.

Tampón endonucleasa S1: NaCl 5 M, AcNa 1,5 M pH4,4, ZnSO₄ 80 mM.

Tampón de parada S1: NH₄Ac 4 M, EDTA 20 mM pH8, 40 µg/ml de ARNt.

Marcaje con Polinucleótido Quinasa T4 y [γ^{32} P]dATP

Técnica empleada para marcar radiativamente la sonda utilizada en el ensayo de protección a endonucleasa S1 de alta resolución, y que se basa en la capacidad de la T4 Polinucleótido Quinasa (PNK) de fosforilar el extremo 5' desfosforilado de un oligonucleótido mediante la incorporación de [γ^{32} P]dATP. La reacción de marcaje se realizó con 50 picomoles de oligonucleótido y 16U de PNK a una temperatura de 37°C durante 45 minutos.

El oligonucleótido marcado se purificó de restos del isótopo radiactivo no incorporado a la sonda, mediante el uso de las columnas de Sephadex Micro Bio Spin P30 Tris Chromatography Columns de Biorad (Ref.: 732-6223). A continuación, se procedió a amplificar por PCR el fragmento deseado como sonda para el ensayo de protección a endonucleasa S1 de alta resolución utilizando dos oligonucleótidos, el marcado con [γ^{32} P]dATP y otro no marcado. Dicha amplificación se purificó en un gel de acrilamida-urea 6% por el método *crush-and-soak* descrito por Sambrook y Russell (2001).

Electroforesis en geles de acrilamida-urea 6%

Estas electroforesis se realizaron tanto para purificar las sondas que se utilizaron en los ensayos de protección a endonucleasa S1 de alta resolución como para analizar los resultados de dichos ensayos.

La preparación de los geles de acrilamida-urea 6% se llevó a cabo siguiendo las indicaciones de Sambrook *et al.*, (1989).

Solución acrilamida-urea 6%: urea 45% (m/v), acrilamida 5,7% (m/v), bisacrilamida 0,3% (m/v), TBE1X.

Tampón de carga para geles de acrilamida: formamida 95%, EDTA 20 mM pH8, azul de bromofenol 0,05% y xilencianol 0,05%.

Ensayos de protección a endonucleasa S1 de baja resolución

Se realizaron siguiendo el método descrito por Hopwood *et al.*, (1985). Esta técnica nos permite detectar la presencia de fragmentos protegidos, pero a diferencia de los ensayos de alta resolución, no nos permite identificar los inicios de transcripción.

En este caso el ARN se hibridó con sondas no marcadas de ADN en las mismas condiciones que para los ensayos de protección a endonucleasa S1 de alta resolución. Los híbridos se analizaron en geles de agarosa alcalinos que posteriormente se transfirieron por *Southern Blot* a una membrana de nylon (Materiales y Métodos, apartado 2.11). Dicha membrana se hibridó frente a una sonda de ADN marcada radiativamente con [α^{32} P]dCTP ó

[$\alpha^{32}\text{P}$]dGTP utilizando el kit Rediprime II Random Prime Labelling System de Amersham Biosciences (Ref.: RPN1633).

Electroforesis en geles alcalinos

Estas electroforesis se utilizaron para analizar los fragmentos protegidos de los ensayos de protección a endonucleasa S1 de baja resolución. Se emplearon geles alcalinos de agarosa 1%, NaCl 50 mM y EDTA 1 mM en un tampón alcalino y utilizando una bomba que permitiese la recircularización del tampón (Hopwood *et al.*, 1985).

Tampón de carga alcalino: 50mM NaOH, EDTA 1 mM, Ficoll 2,5% y orange G dye 0,2%.

Tampón alcalino: NaOH 30 mM, EDTA 1 mM.

2.16. CONSTRUCCIÓN DE UNA GENOTECA DE ADN TOTAL DE *Streptomyces* sp. CO-AC25

Tras digerir parcialmente el ADN total de *S. sp* CO-AC25 con el enzima de restricción *Sau3A* durante un tiempo de incubación determinado empíricamente, se procedió a separar aquellos fragmentos de ADN total de aproximadamente 20 kb. Se realizó un fraccionamiento en gradiente de sacarosa del 10% al 40% en Tris-HCl 25 mM, EDTA 2 mM, NaCl 100 mM, que se centrifugó a 35000 rpm durante 17 horas a 15°C utilizando un rotor SW41Ti. Se recogieron 24 fracciones, de las cuales se cargaron 10 μl en un gel de agarosa 0,7%. Las fracciones cuyo tamaño aproximado era de 15-20 kb y se procedió a ligarlas en el fago λ EMBL-4 previamente digerido con *Bam*HI y *Sall*, enzimas compatibles con *Sau3A* y que evitaban la religación del vector aumentando la eficiencia en la obtención de fagos recombinantes.

Para el empaquetamiento "*in vitro*" de la genoteca se utilizó el kit Packagen Lambda DNA Packaging System de Promega (Ref.: K3154), siguiendo las instrucciones del fabricante. La suspensión de fagos recombinantes en SM (NaCl 0,1 M; MgSO_4 10 mM; Tris-HCl 50 mM pH7,5; gelatina 0,01%) se almacenó a 4°C.

2.17. BIOENSAYOS FRENTE A *Penicillium* Y *Micrococcus*

En placas de LB (Materiales y Métodos, apartado 2.3) se añadió un agar de cobertura SNA con esporas de *Penicillium* ó células de *Micrococcus* sobre el que se colocaron discos de antibiograma impregnados con los extractos cuya actividad se quiso determinar y se incubaron durante toda la noche a 30°C para *Penicillium* y a 37°C en el caso de *Micrococcus*.

2.18. ANÁLISIS POR HPLC

Los análisis por cromatografía líquida de alta presión (HPLC) de los caldos de cultivo de *S. sp.* COAC25 en condiciones analíticas, se llevaron a cabo utilizando una columna Kromasil C8 en fase reversa (100 x 4,6 mm) en gradiente ternario según la tabla 11. Como fases móviles se utilizaron MetOH 100%, NH₄Ac 20 mM pH5, EtOH 20% y Acetonitrilo 10 mM TEA, eluyendo durante 15 minutos a un flujo de 0,7 ml/min, según el gradiente que se detalla a continuación:

	Tiempo (min)	Flujo (ml/min)	%A	%B	%C	%D	Curva
1		0.70	0.0	0.0	80.0	20.0	
2	2.00	0.70	0.0	0.0	80.0	20.0	6
3	4.00	0.70	0.0	0.0	50.0	50.0	3
4	15.00	0.70	100.0	0.0	0.0	0.0	11
5	17.00	0.70	0.0	0.0	80.0	20.0	11

Tabla 11. Gradiente utilizado para el análisis por HPLC.

Los espectros de absorción y la detección del compuesto fungicida se llevaron a cabo mediante un detector de fotodiodeo de array (Waters 996). La extracción de los cromatogramas se realizó a una longitud de onda de 360 nm.

Las muestras analizadas se obtuvieron bien directamente desde caldos de cultivo o desde placas de agar esporuladas en las que primero se fragmenta el agar y posteriormente se realiza una extracción con MetOH 100%.

2.19. EXTRACCIÓN DE METABOLITOS ACTIVOS

Los principios activos producidos por *Streptomyces sp.* CO-AC25 se extrajeron tanto de medio sólido R5 como de medio líquido E1 (Materiales y Métodos, apartado 2.3), el cual contiene en su composición una resina de interacción hidrofóbica XAD-16 que permite la separación de los compuestos bioactivos del resto del caldo de fermentación.

Tras inocular esporas del productor en ambos medios e incubar a 30°C durante 6-7 días en caso del medio sólido o 3-4 días en caso del medio líquido, se procedió a la extracción de los metabolitos activos bien desde el agar con 20 ml de H₂O y 25 ml de acetato de etilo o bien centrifugando el cultivo líquido y extrayendo los metabolitos activos con 45 ml de MetOH 100%. Después de un tiempo en agitación protegiendo de la luz se recoge la fase correspondiente al acetato de etilo o al MetOH y se rotavapora. Obtenido el producto seco se lava con MetOH y se vuelve a rotavaporar disolviendo finalmente en 3-5 ml de MetOH.

Fraccionamiento en columna Kromasil C8 de fase reversa

Una vez obtenida la muestra se procedió a cromatografía y se realizó un fraccionamiento de la misma en condiciones semipreparativas. Para ello, se aplicaron 100 μ l de la muestra resuspendida en MetOH a una columna Kromasil C8 (150 x 10mm) a un flujo de 1,8 ml/min y se le aplicó el gradiente reflejado en la tabla 12.

	Tiempo (min)	Flujo (ml/min)	%A	%B	%C	curva
1	0	1,8	0	80	20	
2	7	1,8	0	80	20	6
3	14	1,8	0	50	50	3
4	35	1,8	100	0	0	11
5	45	1,8	0	20	20	11

Tabla 12. Gradiente utilizado para el fraccionamiento.

2.20. SOPORTE INFORMÁTICO

Para el análisis de secuencias de ADN y aminoácidos se han utilizado los siguientes programas:

DNA Strider (Marck, 1988), *GeneJockey II* y un conjunto de programas denominado *GCG (Genetic Computer Group)* (Deverux *et al.*, 1984). Dentro de *GCG* los programas más usados han sido: *SEQED*, para editar secuencias de ADN; *CODONPREFERENCE*, para la búsqueda de fases abiertas de lectura a partir de una tabla construida a partir de secuencias de 100 genes de *Streptomyces* (Wright y Bibb, 1992); *TRANSLATE*, para la traducción de secuencias de nucleótidos a secuencias de aminoácidos; *BOXSHADE*, para alineamientos de grupos de secuencias.

El programa *BLAST* del *NCBI (National Center for Biotechnology Information)* (Altschul *et al.*, 1997) se ha empleado para comparar secuencias de ADN y de proteínas con secuencias presentes en las bases de datos.

RESULTADOS

3.1. CARACTERIZACIÓN DE METABOLITOS ACTIVOS EN *Streptomyces* sp. CO-AC25

Con el fin de caracterizar los compuestos bioactivos producidos por *Streptomyces* sp. CO-AC25, se procedió a la extracción y caracterización de los mismos. Los caldos de fermentación fueron sometidos a un fraccionamiento con el objeto de localizar e identificar aquellas fracciones en las que se encontraban las actividades biológicas. El fraccionamiento se llevó a cabo en condiciones semipreparativas (Materiales y Métodos, apartado 2.19), se tomaron fracciones de 1,8 ml a intervalos de un minuto y se obtuvieron un total de 35 fracciones. Confirmado el grado de pureza de las mismas, de cada una de ellas se tomaron 500 μ l, que se desecaron y se resuspendieron en 100 μ l de MetOH, con los que se realizaron antibiogramas frente a *Penicillium* y *Micrococcus*, para detectar tanto posibles actividades fungicidas como bactericidas.

El resultado de los antibiogramas se muestra en la figura 13. Se encontraron en el caldo de fermentación de *Streptomyces* sp. CO-AC25 dos actividades separadas en fracciones distintas. Una de las fracciones presentó actividad antifúngica mientras que la otra presentó actividad antibacteriana.

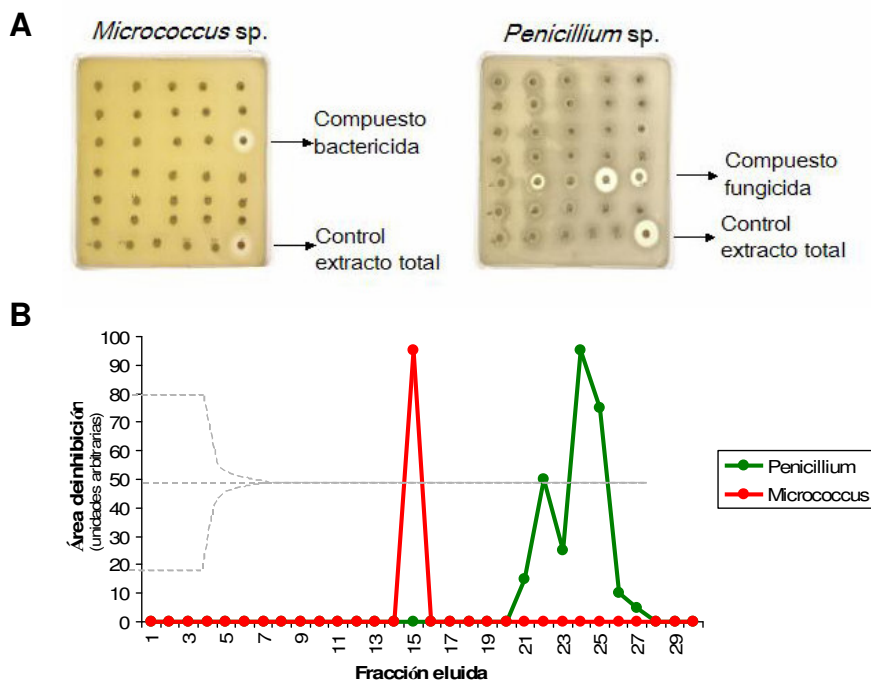


Fig. 13. A. Antibiogramas frente a *Penicillium* y *Micrococcus* de las distintas fracciones obtenidas. **B.** Representación gráfica de las actividades fungicida y bactericida presentes en el caldo de fermentación de *S.* sp. CO-AC25 medidas en unidades arbitrarias.

Resultados

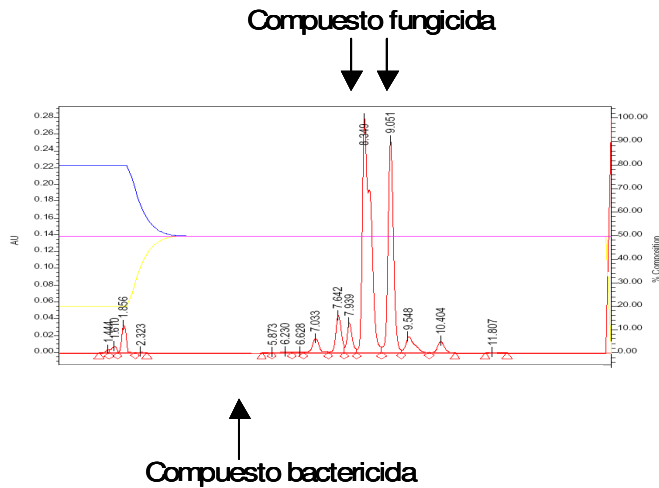


Fig. 14. Cromatograma del caldo de fermentación de *Streptomyces* sp. CO-AC25. Con flechas se señalan los picos correspondientes a las diferentes fracciones obtenidas.

Caracterización de las actividades:

- Actividad bactericida:

Analizando la fracción activa frente a *Micrococcus* no se detectó ningún compuesto que absorbiera en el rango del UV-visible (Figura 14). Por lo tanto, no se tiene ninguna evidencia de cuál puede ser la estructura química del compuesto bactericida producido por *Streptomyces* sp. CO-AC25.

Ensayos de actividad biológica realizados frente a otros dos géneros de bacterias, por un lado frente a una bacteria Gram positiva distinta de *Micrococcus* como es *Streptomyces lividans* TK21 y por otro lado frente a una bacteria Gram negativa como es *E. coli* (Figura 15), permitieron concluir que el compuesto bactericida producido por *Streptomyces* sp. CO-AC25 tiene actividad frente a bacterias Gram positivas pero no parece ser activo frente a Gram negativas.

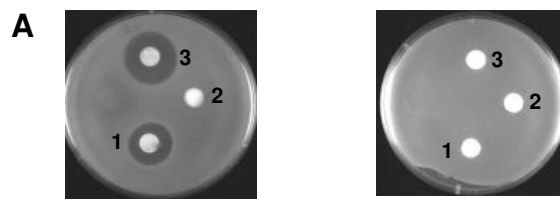


Fig. 15. Ensayos de actividad biológica frente a *S. sp. TK21* (A) y frente a *E. coli* (B) del compuesto bactericida (1) y fungicida (2) producidos por *Streptomyces* sp. CO-AC25. (3): caldo de fermentación.

- Actividad fungicida:

Analizando las fracciones activas frente a *Penicillium* se detectó la presencia de compuestos que absorbían fuertemente en el ultravioleta y cuyo espectro de absorción era característico de macrólidos polienos del grupo de los oxopentaenos, con un máximo de absorción a 369 nm (Figura 16).

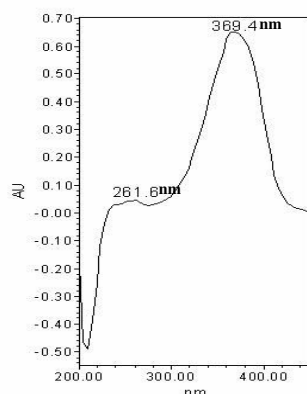


Fig. 16. Espectro ultravioleta característico de macrólidos oxopentaenos.

El cromatograma que se obtiene tras analizar mediante HPLC el caldo de fermentación de *S. sp.* CO-AC25 (Figura 14) evidenció la presencia de dos compuestos con tiempos de retención próximos y espectros de absorción muy similares, que se correspondían con los halos de inhibición observados en el antibiograma realizado frente a *Penicillium* (Figura 13.A). La proporción relativa de ambos compuestos variaba de unas preparaciones a otras, por ello se procedió a determinar si realmente se trataba de dos compuestos diferentes o eran realmente la transformación de uno en otro debido a inestabilidad. En ensayos posteriores se modificaron las condiciones de la extracción, realizándose éstas fundamentalmente en oscuridad. Así se pudo conseguir una preparación donde el compuesto de menor movilidad cromatográfica representaba el 95% del total (Figura 17). Por tanto, parece que el pico de mayor movilidad cromatográfica procede de una degradación del compuesto de menor movilidad por efecto de la luz ultravioleta.

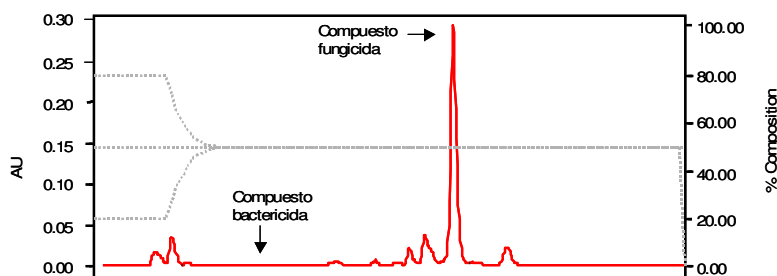


Fig. 17. Cromatograma del caldo de fermentación de *S. sp.* CO-AC25 extraído en condiciones de oscuridad total.

Determinación de la estructura química del compuesto fungicida:

La determinación estructural del compuesto fungicida se realizó en colaboración con Galilaeus OY (Turku, Finlandia). Los espectros de masas y experimentos de Resonancia Magnética Nuclear (RMN) determinaron que dicho compuesto se corresponde con flavofungina I a la que se le atribuye una masa molecular de 650 (R. Bognár *et al.*, 1970) (Figura 18).

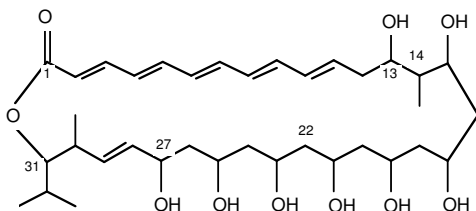


Fig. 18. Estructura química de flavofungina I.

Se trata por tanto de un compuesto oxopentaeno como sugería el espectro UV-visible. Debido a la falta de información sobre este tipo de compuestos, se consideró interesante disponer de genes biosintéticos que en un futuro permitieran manipulaciones genéticas.

3.2. ANÁLISIS DE SISTEMAS DE MANIPULACIÓN GENÉTICA

Con el fin de explorar la disponibilidad de herramientas genéticas para manipular *Streptomyces* sp. CO-AC25, se ensayaron distintas metodologías para la introducción de ADN y, de forma paralela, se analizó la sensibilidad de la cepa frente a diferentes antibióticos.

Los ensayos de sensibilidad a antibióticos se realizaron en medio sólido R5 (Materiales y Métodos, apartado 2.3). Se comprobó que la cepa era resistente a eritromicina, neomicina, kanamicina e higromicina y sensible a apramicina, gentamicina y thioestrepton a las concentraciones indicadas en Materiales y Métodos (apartado 2.4). En el caso del antibiótico thioestrepton se observó siempre un fondo de resistencia, que llevó a descartarlo como marcador de selección.

Para la introducción de ADN se ensayó la transformación con plásmidos y/o cósmidos mediante la formación de protoplastos, la transfección, la electroporación y la conjugación, tanto entre distintas especies de *Streptomyces* (conjugación intraespecífica) como entre *Streptomyces* y *E. coli* (conjugación interespecífica).

- Susceptibilidad para formar protoplastos: se siguieron los protocolos utilizados para otras especies de *Streptomyces* descritos en Hopwood *et al.*, (1985) y no se obtuvieron resultados satisfactorios, por lo que se introdujeron modificaciones de dichos protocolos, ya que los parámetros óptimos para la formación de protoplastos varían de unas cepas a otras. Estas variaciones fueron:

- el medio elegido para la obtención del micelio así como la concentración de glicina adicionada al medio. Se ensayaron los medios líquidos Yeme, TSB, SYM2, E1, R5 y YES (Materiales y Métodos, apartado 2.3) a concentraciones de glicina entre 0,3%-0,5%.

- el tiempo de incubación del cultivo para la obtención del micelio usado para la preparación de los protoplastos. Se probaron tiempos de incubación de 24 y 48 horas a 30°C.

- la concentración de lisozima empleada para la lisis del micelio y posterior formación de los protoplastos.

Ninguna de las modificaciones realizadas permitió la obtención de protoplastos de la cepa por lo que se descartó la transformación como método de manipulación genética.

- **Susceptibilidad de *S. sp.* CO-AC25 a actinofagos:** se ensayaron actinofagos derivados de ϕ C31, tanto att⁺ (como es el fago KC212; integración específica) como att⁻ (fago PM1; integración facilitada por insertos).

Se infectaron con ambos tipos de fagos esporas de *Streptomyces sp.* CO-AC25, y en ambos casos se obtuvieron transductantes. No obstante, resultó sorprendente que para el fago PM1 se obtuvieron lisógenos aun cuando éste no llevaba clonado ningún fragmento de ADN homólogo al cromosoma. Ante la posibilidad que se integrara vía genes de resistencia a higromicina o thiostrepton, se sustituyeron ambos por el gen de resistencia a apramicina y los resultados obtenidos en este caso fueron idénticos. La posibilidad de la presencia de un trasposón en el fago se descartó mediante el análisis de fragmentos de restricción por *Southern Blot*. Así mismo mediante el tratamiento con pirofosfato se descartó la posibilidad de que el fago estuviera directamente asociado a las paredes celulares. Ante lo confuso de los resultados se rechazó el uso de estos vectores como herramientas para introducir ADN en *Streptomyces sp.* CO-AC25.

- **Electroporación:** se prepararon células electrocompetentes de *S. sp.* CO-AC25, creciendo previamente dicha cepa en medio líquido CRM (Materiales y Métodos, apartado 2.3), y se siguió el protocolo descrito por Pigac y Schrepf, 1995. El electroporador empleado fue el Micropulser de Biorad y se utilizó siguiendo las instrucciones del fabricante. Dicho método tampoco proporcionó resultados satisfactorios.

- **Conjugación:** a consecuencia de los resultados negativos obtenidos al intentar desarrollar un sistema de transformación para *Streptomyces sp.* CO-AC25 con las técnicas anteriormente descritas, se pensó como alternativa en la conjugación del mismo con otros microorganismos.

Los ensayos se realizaron con dos tipos de plásmidos. Por un lado, se conjugó la cepa *Streptomyces sp.* CO-AC25 con un plásmido integrativo como es el pSET152 (Bierman *et al.*, 1992) y, por otro, con un plásmido suicida, como es el pOJ260 (Bierman *et al.*, 1992). Ambos plásmidos poseen el origen de transferencia (*oriT*) del plásmido RK2, el origen de replicación de *E. coli*, un sitio de clonación múltiple dentro del *lacZ α* y un marcador de selección operativo tanto para *E. coli* como para *Streptomyces*, la apramicina. El pSET152 cuenta con la ventaja de tener incorporada la maquinaria de integración del actinofago ϕ C31 mientras que el pOJ260 es

un plásmido suicida que recombina por el fragmento homólogo de ADN de *Streptomyces* que lleva clonado. Se emplearon distintas cepas donadoras de *E. coli* para llevar a cabo dicho ensayo, por un lado se utilizaron las células S17-1 (Simon *et al.*, 1983) y por otro, las células ET12567 (pUZ8002) (Paget *et al.*, 1999), que poseen un sistema deficiente en metilación.

Tras varias modificaciones del protocolo propuesto por Mazodier *et al.*, (1989) se estableció el protocolo que permitió la obtención de exconjugantes y se vio que la eficacia de conjugación era elevada con el plásmido pSET152 (Materiales y Métodos, apartado 2.8). La conjugación ofrecía, por tanto, una excelente herramienta para la manipulación genética de *Streptomyces* sp. CO-AC25.

3.3. AISLAMIENTO Y CARACTERIZACIÓN DE GENES BIOSINTÉTICOS DE LOS COMPUESTOS BIOACTIVOS

Una vez determinada la naturaleza química del compuesto fungicida como un macrólido polieno, se procedió al aislamiento de sus genes biosintéticos. Para ello se generó una librería genómica de *S. sp.* CO-AC25 en el fago λ EMBL-4 tal y como se indica en Materiales y Métodos, apartado 2.16. La titulación de esta genoteca resultó ser de $7 \cdot 10^5$ ufp/ml. Para el aislamiento de fagos recombinantes por hibridación "*in situ*" se utilizaron 10^4 ufp repartidas en 12 placas.

Escrutinio de la genoteca de *Streptomyces* sp. CO-AC25

Para el rastreo de la genoteca se generaron dos sondas:

(a) fragmentos amplificados por PCR del cromosoma de *Streptomyces* sp. CO-AC25, utilizando unos oligonucleótidos diseñados para amplificar fragmentos internos de dominios ketosintasa (KS), *PimR* y *PimD* (Seco *et al.*, 2004) que habían proporcionado buenos resultados en otros sistemas.

(b) un fragmento heterólogo procedente del *cluster* de biosíntesis de los tetraenos CE-108 y rimocidina producidos por *Streptomyces diastaticus* var. 108 (Seco *et al.*, 2004), codificante de RimA.

Mientras que los fragmentos obtenidos por PCR con los oligonucleótidos *PimR* y *PimD* no se llegaron a utilizar como sondas para escrutar la genoteca, a partir de la sonda heteróloga se aislaron distintas regiones del cromosoma que presentaban un alto grado de homología a PKSs Tipo I.

Aislamiento de genes biosintéticos a partir de la sonda heteróloga

La sonda heteróloga utilizada para escrutar la genoteca fue obtenida del plásmido pSM23.8. Fragmento *SacI* de 3,1 kb clonado en el vector pIJ2925, que codifica la PKS del módulo de carga de rimocidina (RimA) producida por *Streptomyces diastaticus* var.108 (Seco *et al.*, 2004) (Figura 19) y que se marcó mediante *Nick Translation* según se describe en Materiales y Métodos, apartado 2.12.

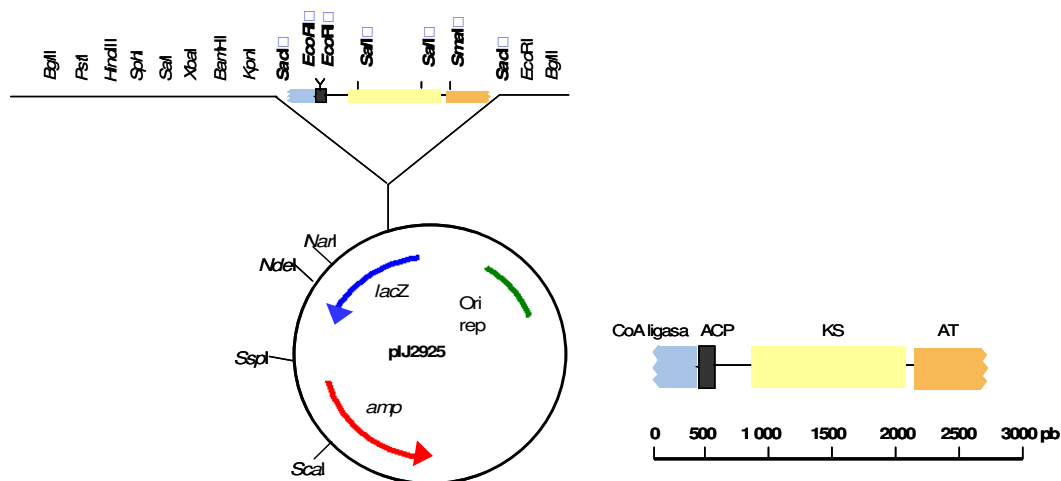


Fig. 19. Plásmido pSM23.8 (Seco *et al.*, 2004), cuyo fragmento *SacI* fue usado como sonda para rastrear la genoteca de *S. sp.* CO-AC25.

Se aislaron tres regiones independientes del genoma del productor en tres fagos recombinantes. Los análisis de hibridación confirmaron la homología de los fragmentos aislados con la sonda empleada, así como la integridad de los mismos respecto al genoma y la inexistencia de reorganizaciones génicas. Los análisis tanto de restricción como de homologías por *Southern Blot* también evidenciaron que las regiones aisladas del genoma no eran solapantes entre sí.

Los fagos recombinantes aislados recibieron los nombres de EMBL-4/11J, EMBL-4/1B y EMBL-4/P1 (Figura 20), y contienen regiones del cromosoma de aproximadamente 14 kb, 15 kb y 13 kb respectivamente. Dichos fagos fueron mapeados con diferentes enzimas de restricción, y las regiones altamente homólogas se subclonaron y secuenciaron evidenciando un alto grado de homología con PKSs Tipo I.

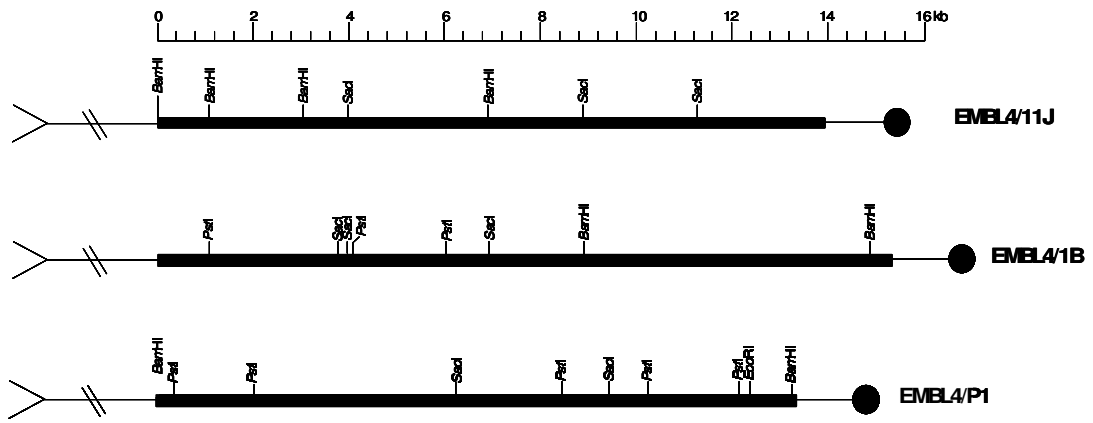


Fig. 20. Mapa de restricción de las regiones aisladas de la genoteca de *S. sp.* CO-AC25 a partir de la sonda heteróloga.

Estudios de interrupción génica

Para determinar la pertenencia de las regiones clonadas al *cluster* de flavofungina I, se procedió a realizar experimentos de interrupción génica mediante inserción inactivante utilizando plásmidos suicidas y la conjugación con *E. coli* como sistema para introducir el ADN.

Fago EMBL-4/11J:

Se procedió a secuenciar parcialmente el fragmento interno de dicho fago que hibridaba con la sonda utilizada para rastrear la genoteca. Se trataba de un fragmento de ADN *SacI*-*Bam*HI de 3,1 kb (nucleótidos 4952-8159 del Anexo I) que fue clonado en el vector pJ2925 y secuenciado íntegramente. Dicha construcción recibió el nombre de pAS11J2-1 y se comprobó que se trataba de un fragmento interno a un gen que codificaba una proteína con similitud a PKSs Tipo I lo que determinó su idoneidad para la interrupción génica. Se procedió a su clonaje en forma de *Eco*RI/*Hind*III en el plásmido suicida y conjugativo pOJ260 (Bierman *et al.*, 1992) y la construcción resultante recibió el nombre de pLM31 (Figura 21).

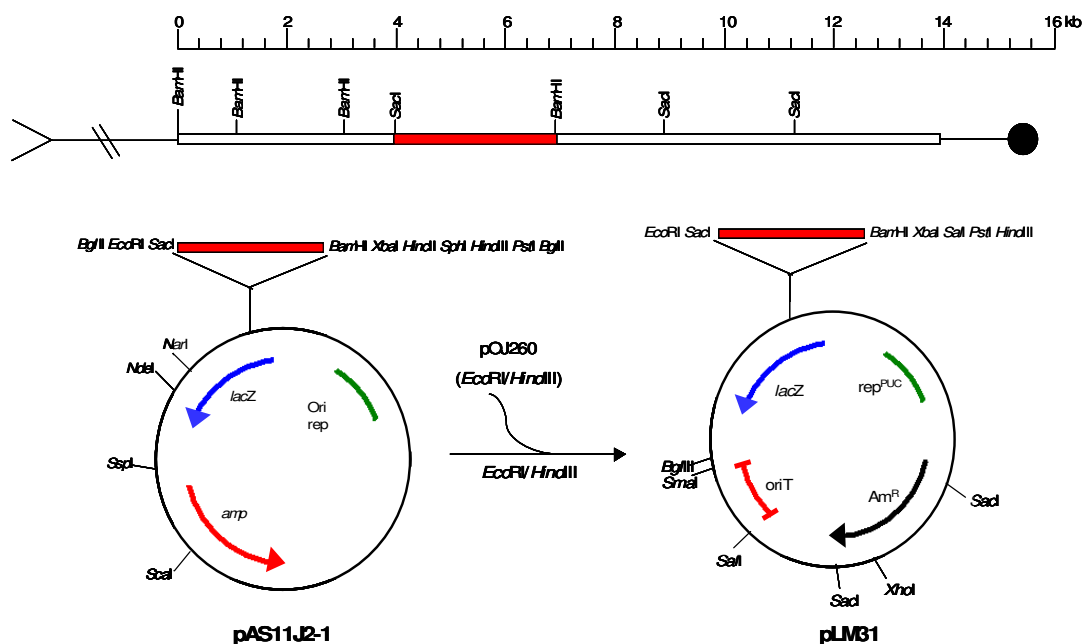


Fig. 21. Esquema de la construcción del plásmido pLM31 empleado en la conjugación con el objeto de obtener la interrupción génica.

Con dicho plásmido se llevó a cabo la conjugación interespecífica *E. coli* S17-1/*Streptomyces* sp. CO-AC25 (Materiales y Métodos, apartado 2.8) y se obtuvo el exconjugante CO-AC25::pLM31, con una inserción en el cromosoma que interrumpía el gen correspondiente al fragmento clonado (Figura 22). La correcta inserción se confirmó mediante hibridación por *Southern Blot* usando como sonda el fragmento *SacI*-*Bam*HI de 3,1 kb.

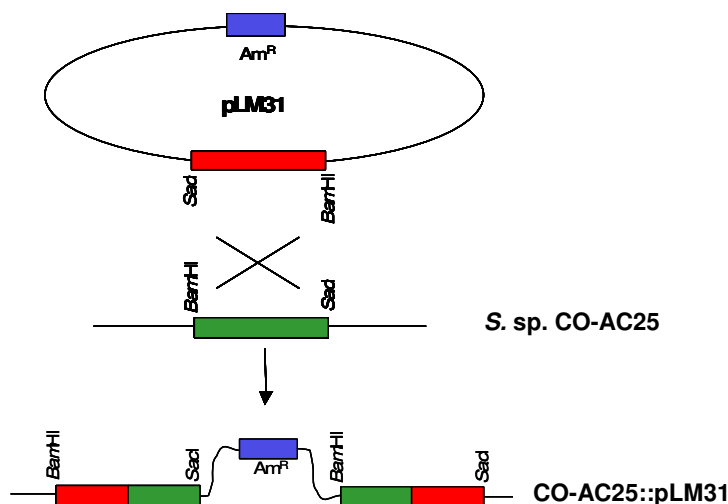


Fig. 22. Esquema de la interrupción inactivante en la región perteneciente al fago recombinante EMBL-4/11J.

Fago EMBL-4/1B:

La interrupción génica se realizó de forma similar que en el caso anterior. Se procedió al clonaje y secuenciación del fragmento interno del fago que hibridaba con la sonda empleada para rastrear la genoteca, se trataba de un fragmento *SacI* de 2,9 kb (nucleótidos 7607-10543 del Anexo II) que se clonó en el vector pIJ2925. A esta construcción se le dio el nombre de pFPZ12. El fragmento clonado se secuenció y se comprobó que se trataba de un fragmento interno a un gen que codificaba una proteína que presentaba alta homología con PKSs Tipo I. El fragmento de ADN *SacI* de 2,9 kb de pFPZ12 se clonó en el plásmido suicida y conjugativo pOJ260 (Bierman *et al.*, 1992) como *EcoRI*/*Hind*III y el plásmido originado se denominó pLM32 (Figura 23). Con dicho plásmido se procedió a la conjugación interespecífica y se obtuvo el exconjugante CO-AC25::pLM32 (Figura 24). La inserción se confirmó mediante hibridación del ADN total del exconjugante CO-AC25::pLM32 con el fragmento *SacI* de 2,9 kb utilizado para llevar a cabo la interrupción y que se usó como sonda.

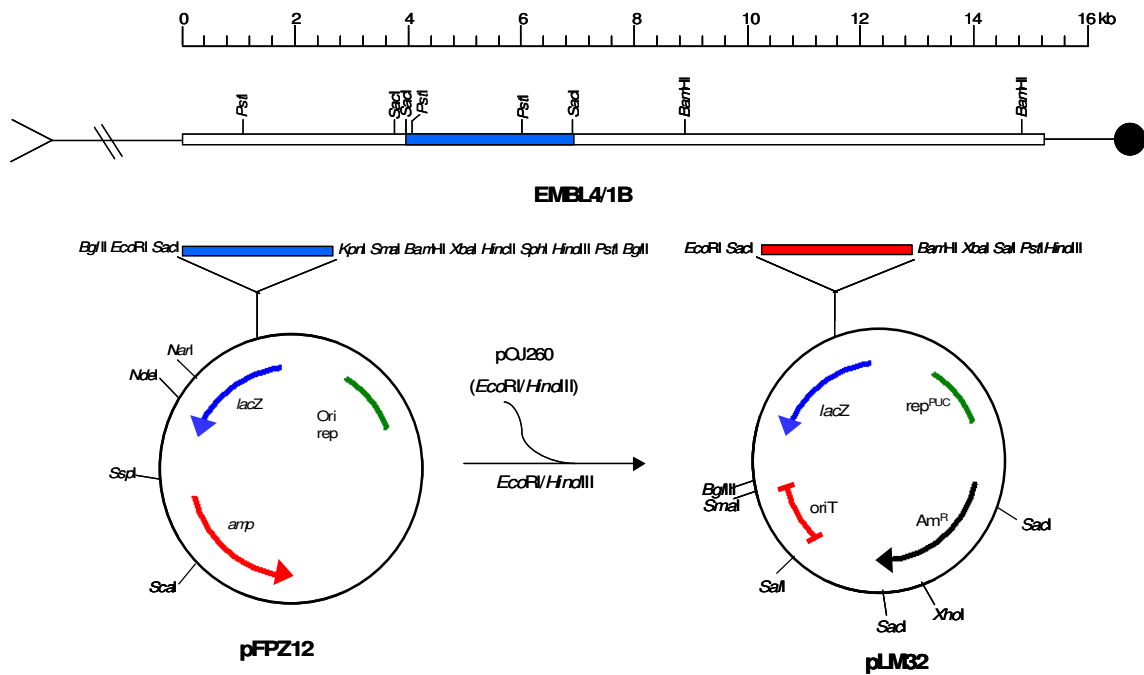


Fig. 23. Esquema de la construcción del plásmido pLM32 utilizado en la interrupción génica.

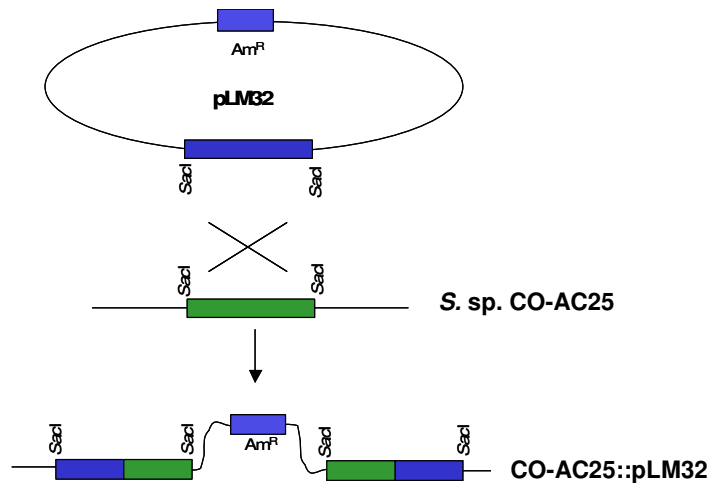


Fig. 24. Esquema de la interrupción inactivante en la región clonada en el fago recombinante EMBL4/1B.

Fago EMBL-4/P1:

Como en los casos anteriores, se clonó en el vector pJ2925 el fragmento interno *PstI* de 4,5 kb proveniente del plásmido pTC4 y que hibridaba con la sonda usada para rastrear la genoteca dando origen al plásmido pTC8, tras ser secuenciado y evidenciar su homología con PKSs Tipo I se clonó en el plásmido pOJ260 (Bierman *et al.*, 1992). La construcción resultante, pLM33 (Figura 25) se utilizó para la conjugación interespecífica y el exconjugante obtenido se llamó CO-AC25::pLM33 (Figura 26). Mediante *Southern Blot* usando como sonda el fragmento *PstI* de 4,5 kb se confirmó que la inserción producida era la esperada.

Resultados

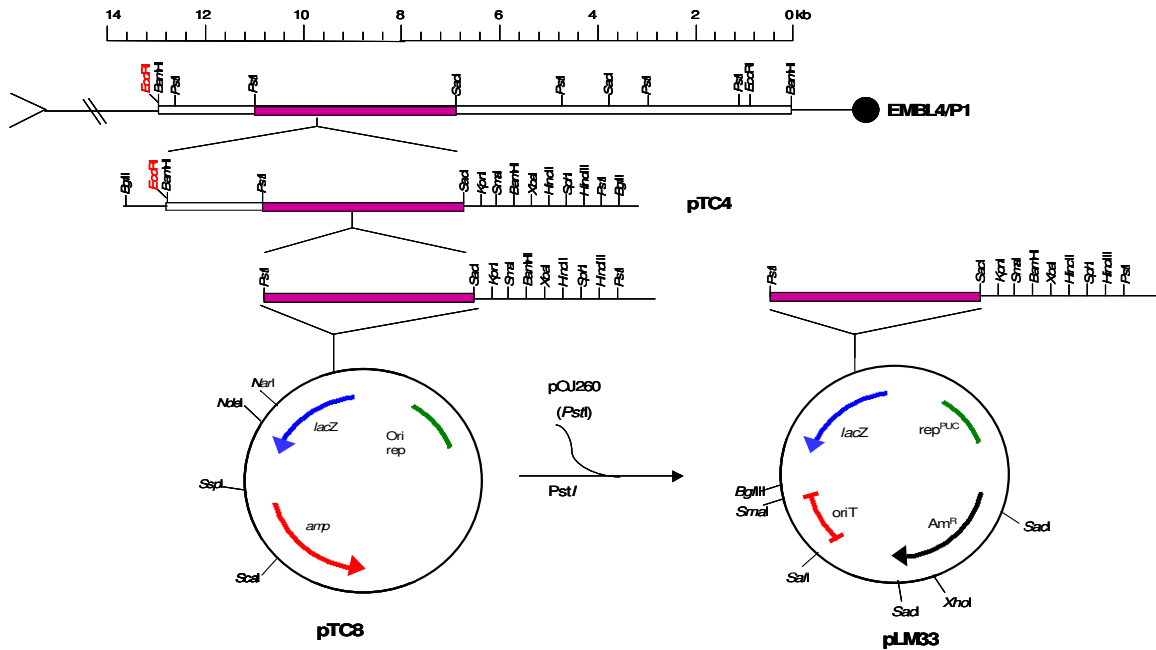


Fig. 25. Esquema de la construcción del plásmido pTC8. En rojo se señala la diana de restricción *EcoRI*, la cual no pertenece al cromosoma de *S. sp.* CO-AC25 sino al fago EMBL4/P1.

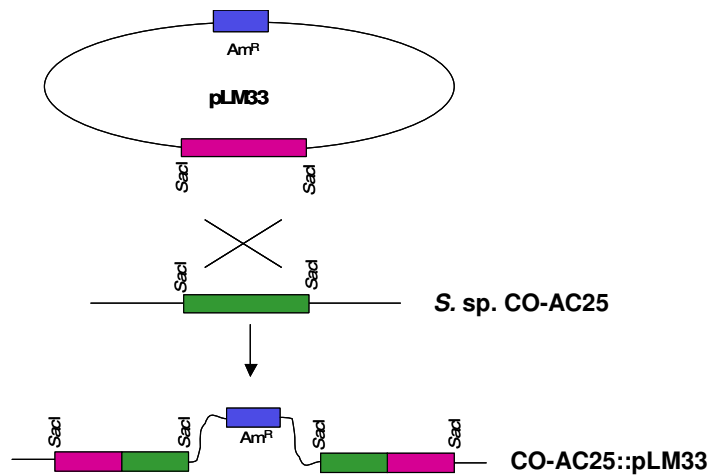


Fig. 26. Esquema de la disrupción inactivante para el exconjugante *S. sp.* CO-AC25:: pLM33.

En resumen, las tres disrupciones génicas llevadas a cabo, tuvieron lugar en los genes correspondientes a los genes de las PKSs que codificaban, encontrándose en cada caso los cambios esperados en los patrones de restricción al ser comparados con la cepa silvestre y sugiriendo ausencia de reordenaciones cromosómicas. Como ejemplo en la figura 27.A se ilustra el esquema de integración del plásmido pLM31 en la región cromosómica de *S. sp.* CO-AC25, de cuyo resultado se generan una banda *Bam*HI y una banda *Pst*I inexistentes en la cepa silvestre. En la figura 27.B se muestra el *Southern Blot* que confirma dicha integración.

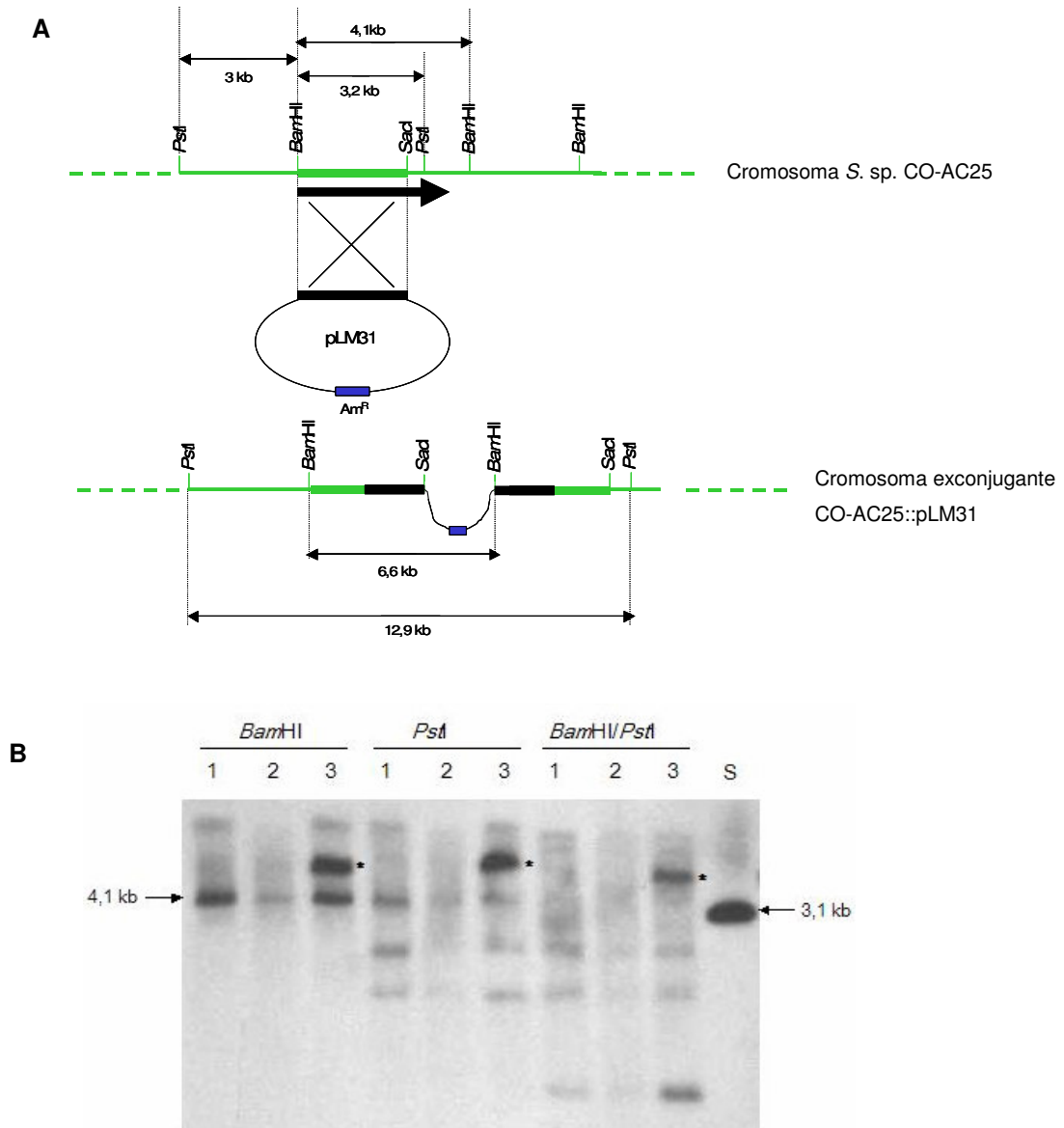


Fig. 27. A. Esquema de integración del plásmido pLM31 en la región cromosómica de *S. sp.* CO-AC25. **B.** *Southern Blot* de confirmación: (1) *S. sp.* CO-AC25; (2) CO-AC25::pSET152 (control de la conjugación); (3) CO-AC25::pLM31. Con asterisco (*) se señalan las bandas *Bam*HI y *Pst*I generadas; (S) fragmento *Bam*HI/*Sac*I de 3,1 kb.

Análisis de la actividad biocida de los exconjugantes

Una vez confirmadas las interrupciones génicas de los tres exconjugantes mediante *Southern Blot*, se procedió a evaluar la capacidad de los mismos para producir el oxopentaeno flavofungina I, mediante la realización de antibiogramas y HPLC, tal y como se describe en Materiales y Métodos, apartados 2.17 y 2.18.

Los patrones de producción mediante HPLC de los tres exconjugantes obtenidos confirmaron que mientras que los exconjugantes CO-AC25::pLM31 y CO-AC25::pLM32 mostraban un fenotipo no productor de flavofungina I, para el exconjugante CO-AC25::pLM33 la producción de flavofungina I no sólo no se había interrumpido sino que era mayor que la de la cepa silvestre (Figura 28).

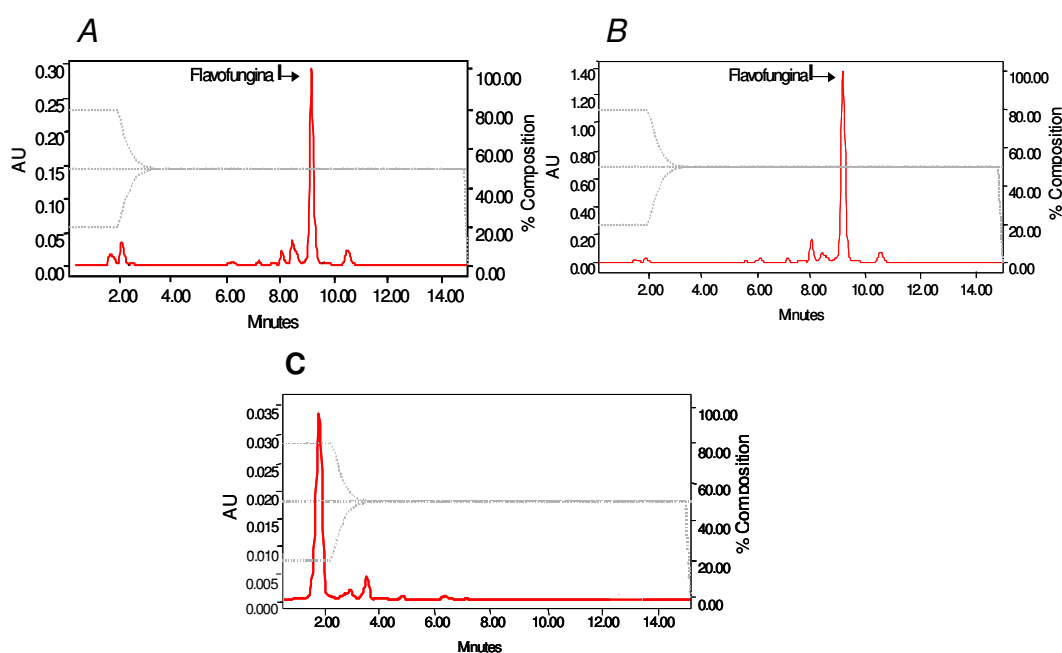


Fig. 28. Análisis de producción de la cepa silvestre CO-AC25::pSET152 (A); el exconjugante CO-AC25::pLM33 (B), dónde la producción de flavofungina I está sobreexpresada; y el exconjugante CO-AC25::pLM31 (C), dónde existe ausencia del pico de producción correspondiente a flavofungina I.

La realización de los correspondientes antibiogramas frente a *Penicillium* y frente a *Micrococcus* pusieron de manifiesto la ausencia de actividad fungicida pero no bactericida en los caldos de fermentación de los exconjugantes CO-AC25::pLM31 y CO-AC25::pLM32 (Figura 29); sin embargo, en el caso del exconjugante CO-AC25::pLM33, los antibiogramas realizados frente a *Micrococcus* y frente a *Streptomyces lividans* TK21 pusieron de manifiesto la ausencia de actividad bactericida en su caldo de fermentación (Figura 30).

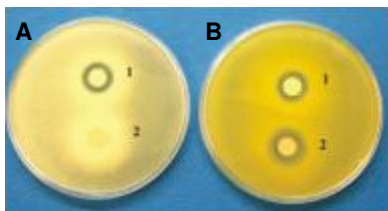


Fig. 29. Bioensayo para la detección de actividad fungicida frente a *Penicillium* (A) y frente a *Micrococcus* (B) del exconjugante CO-AC25::pLM31 (2) y de la cepa silvestre (1).

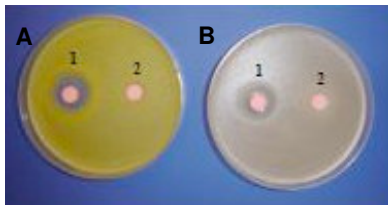


Fig. 30. Bioensayo para la detección de actividad bactericida frente a *Micrococcus* (A) y frente a *S. lividans* TK21 (B) del exconjugante CO-AC25::pLM33 (2) y de la cepa silvestre (1).

En base a estos resultados, se puede concluir que las regiones génicas del cromosoma de *Streptomyces* sp. CO-AC25 clonadas en los fagos recombinantes EMBL4/11J y EMBL4/1B están implicadas en la biosíntesis del macrólido oxopentaeno flavofungina I, y que ambas forman parte del mismo *cluster*. Por otro lado la región génica clonada en el fago EMBL4/P1 interviene en la biosíntesis del compuesto bactericida producido por *Streptomyces* sp. CO-AC25. Se ha observado un aumento de producción del compuesto fungicida cuando la ruta de biosíntesis del compuesto bactericida estaba interrumpida y viceversa.

Ampliación de la región clonada del *cluster* de flavofungina I

A partir de los fagos aislados y mediante la técnica de *chromosome walking* se procedió al aislamiento de fagos solapantes a partir de los ya aislados.

Para avanzar en el cromosoma corriente abajo del fago recombinante EMBL4/11J, se seleccionó como sonda un fragmento de 2,2 kb *SacI/EcoRI* proveniente del plásmido pAS13 situado en el extremo de dicho fago. Así se obtuvo un nuevo fago solapante con el anterior, el fago EMBL4/Db, que abarcaba una región cromosómica de 14,3 kb, de las cuales 8,7 kb solapaban con el fago EMBL4/11J (Figura 31.A (a)).

El proceso se repitió con un fragmento *SphI* de 2 kb aislado del plásmido pG45 situado en el extremo distal del fago EMBL4/Db. Se aisló un nuevo fago solapante al que se llamó EMBL4/b, su tamaño era de 19,2 kb y nos permitía avanzar 10,8 kb corriente abajo en el cromosoma de *Streptomyces* sp. CO-AC25 (Figura 31.A (b)).

Por otro lado y para aislar un fago solapante corriente arriba del fago EMBL4/11J, se utilizó como sonda un fragmento *BamHI* de 2 kb proveniente del plásmido pAS11J7, así se aisló el fago recombinante EMBL4/F1 que aportaba 9,5 kb corriente arriba a toda la región previamente aislada (Figura 31.A (c)).

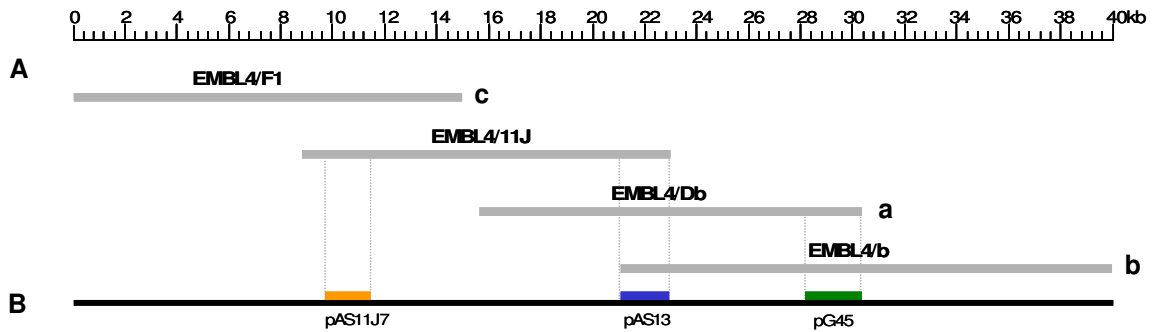


Fig. 31. A. Fagos solapantes obtenidos a partir del rastreo de la genoteca de *S. sp.* CO-AC25.

B. Región del cromosoma de *S. sp.* CO-AC25 comprendida en los fagos solapantes aislados. Con colores se marcan las sondas utilizadas en el rastreo de la genoteca.

Con el fin de avanzar corriente arriba del fago recombinante EMBL4/1B, se procedió una vez más a escrutar la librería genómica de *Streptomyces sp.* CO-AC25. Como sonda se utilizó un fragmento *Bam*HI/*Eco*RI del plásmido pFPZ15, con un tamaño de 600 pb y que permitió aislar un fago solapante con el EMBL4/1B al que se le dio el nombre de EMBL4/K. Dicho fago tenía un tamaño de 14,4 kb y proporcionaba un fragmento nuevo corriente arriba de 6,5 kb (Figura 32.A (a)).

Utilizando como sonda un fragmento de ADN del fago EMBL4/K (fragmento de 400 pb *Bg*II/*Eco*RI del plásmido pK2.1) se aisló un nuevo fago solapante con los dos anteriores. El fago EMBL4/G10 tenía un tamaño de 19,8 kb y permitía avanzar 4,8 kb corriente arriba en el cromosoma de *S. sp.* CO-AC25 (Figura 32.A (b)).

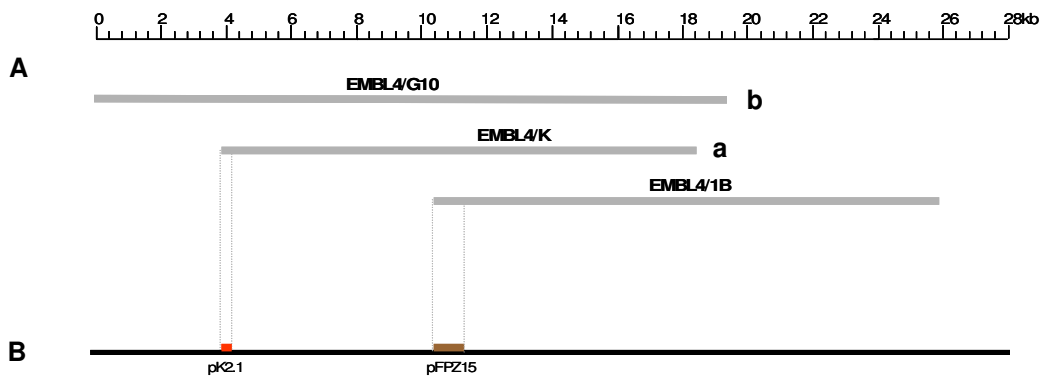


Fig. 32. A. Fagos solapantes obtenidos a partir del rastreo de la genoteca de *S. sp.* CO-AC25. **B.**

Región cromosómica de *S. sp.* CO-AC25 aislada a partir de dichos fagos. Con colores se marcan las sondas utilizadas en el rastreo.

Mediante experimentos de *Southern Blot* se constató que dichos fagos representaban la estructura del genoma en las regiones aisladas y que no existían reorganizaciones génicas. Una vez hecho esto, se procedió a un mapeo de los fagos aislados con enzimas de restricción tales como *Bam*HI, *Eco*RI, *Sac*I, *Bgl*II, *Pst*I, *Xho*I, *Sph*I y al subclonaje de los fragmentos obtenidos en los vectores pIJ2925, pUC18 y pSU20 para su posterior secuenciación.

El mapeo del fago recombinante EMBL4/G10 puso de manifiesto la presencia de un fragmento *Sac*I de tamaño similar a otro fragmento *Sac*I interno del fago EMBL4/b. La secuencia de ambos fragmentos resultó ser idéntica, lo cual implicaba que los fagos recombinantes EMBL4/G10 y EMBL4/b eran solapantes y que se había conseguido aislar una región continua del cromosoma de *Streptomyces* sp. CO-AC25 que abarcaba una totalidad de 62 kb (Figura 33).

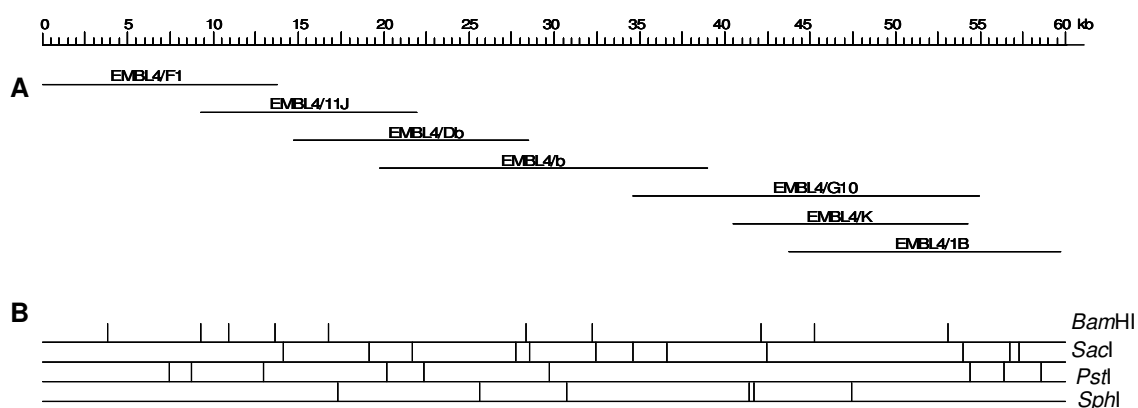


Fig. 33. A. Disposición de los fagos recombinantes aislados de la genoteca de *S. sp.* CO-AC25. **B.** Mapa de restricción de la región cromosómica aislada con las enzimas *Bam*HI, *Sac*I, *Pst*I y *Sph*I.

Con la intención de caracterizar algo más el *cluster* biosintético del compuesto bactericida se llevó a cabo la técnica de *chromosome walking* en el fago recombinante EMBL4/P1. Se aisló un nuevo fago recombinante al que se llamó EMBL4/ABC. Para su aislamiento se utilizó como sonda un fragmento *Eco*RI/*Bam*HI de 750 pb obtenido a partir del plásmido pTC2. El fago EMBL4/ABC aislado tenía un tamaño de 14 kb y nos permitía avanzar 11 kb corriente abajo del fago EMBL4/P1. De igual forma que en los casos anteriores, ambos fagos fueron mapeados con diferentes enzimas de restricción y mediante *Southern Blot* se comprobó que no existían reorganizaciones génicas y que la estructura de estos fagos era la misma que la que presentaba el cromosoma de *Streptomyces* sp. CO-AC25 (Figura 34).

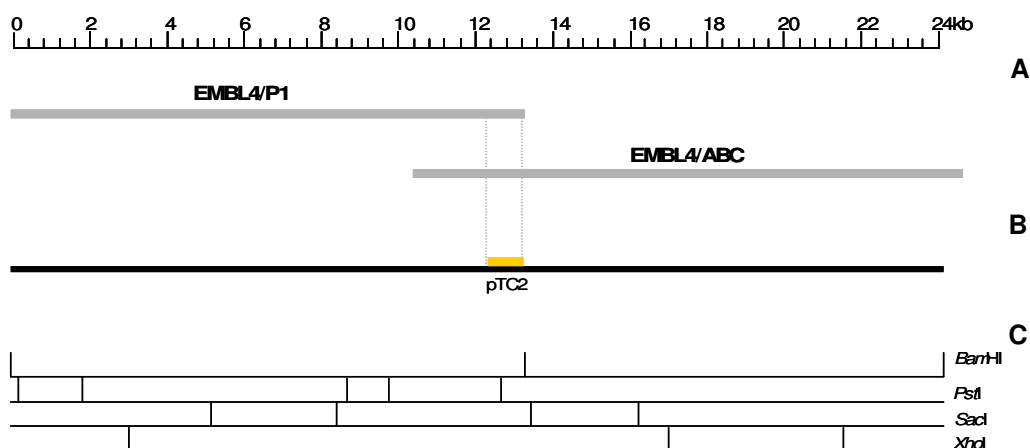


Fig. 34. A. Disposición de los fagos solapantes EMBL4/P1 y EMBL4/ABC. B. Región cromosómica aislada de *Streptomyces* sp. CO-AC25 contenida en dichos fagos. En amarillo se marca la sonda utilizada para rastrear la librería genómica. C. Mapa de restricción de ambos fagos con las enzimas *Bam*HI, *Pst*I, *Sac*I y *Xho*I.

3.4. SECUENCIACIÓN Y DETERMINACIÓN DE REGIONES ABIERTAS DE LECTURA DE LA PKS DE FLAVOFUNGINA I

Los experimentos descritos anteriormente llevaron a disponer de dos regiones cromosómicas no solapantes implicadas en la biosíntesis de flavofungina I. Por un lado un fragmento de 35323 pb (Anexo I) y por otro un fragmento de 13955 pb (Anexo II), ambos separados entre sí por un fragmento de ADN no secuenciado que se estima en aproximadamente 3200 pb. De los análisis se deduce la presencia de seis regiones abiertas de lectura, cuatro de ellas completas y dos incompletas. De los seis genes deducidos, todos menos uno se transcriben en la misma dirección (Figura 35).

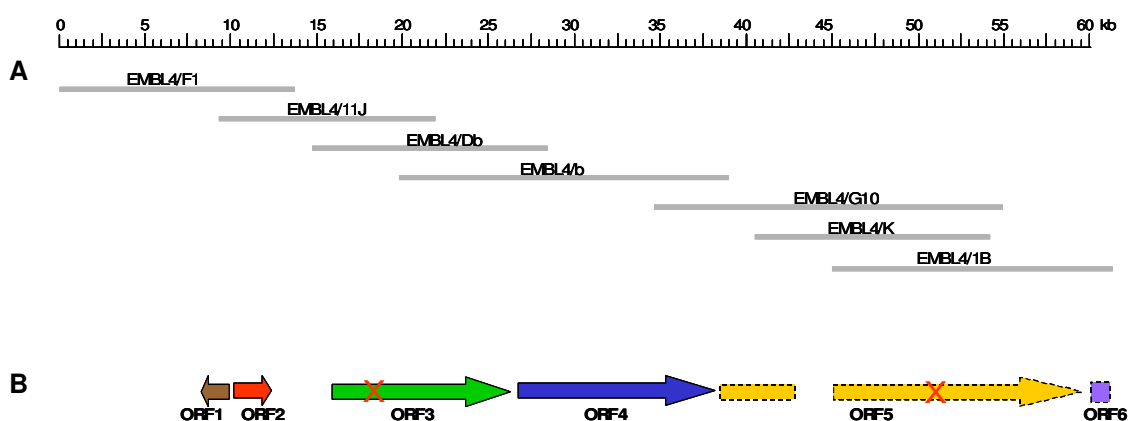


Fig. 35. A. Disposición de los fagos recombinantes que abarcan una región del cromosoma de *Streptomyces* sp. CO-AC25 algo mayor de 60 kb. B. Genes deducidos de la región clonada. Con una cruz roja se marcan los puntos dónde se han llevado a cabo las interrupciones génicas. Con líneas de puntos aparecen las ORFs que se encuentran incompletas.

Las seis regiones abiertas de lectura fueron analizadas:

ORF1: comienza con un codón de inicio ATG (metionina) en la posición 938 precedido por un probable sitio de unión al ribosoma (RBS) GAAAAGAGG, 12 nucleótidos corriente arriba del codón de inicio (Anexo I). Termina en la posición 363 con un codón de parada traduccional TGA. Se trata de un gen de 575 pb con un contenido en G+C del 71,3%. Este gen es el único que se transcribe en la dirección opuesta al resto de los genes analizados en este trabajo.

ORF2: comienza con un codón de inicio ATG en la posición 1119, 3 nucleótidos corriente abajo de un probable sitio de unión al ribosoma AAGGAGAA. Se trata de un gen de 969 pb que termina en un codón de terminación TAG en la posición 2087(Anexo I). Su contenido en G+C es de 75,3%.

ORF3: comienza con un codón de inicio ATG en la posición 3872, precedido de un probable sitio de unión al ribosoma 5 nucleótidos corriente arriba AGAGG. Termina en la posición 16987 con el codón de terminación TGA (Anexo I). Se trata de un gen que comprende 13116 pb y con un contenido en G+C del 76%. Dicho gen recibe el nombre de *flvA*.

ORF4: comienza en la posición 17204 con un codón de inicio TGT (valina), con un probable sitio de unión al ribosoma GAGGA 2 nucleótidos corriente arriba de dicho inicio. Comprende un gen de 14209 pb, con un contenido aproximado en G+C del 73%, que termina en el codón de terminación TGA en posición 31412 (Anexo I). Se trata del gen *flvB*.

ORF5: Recibe el nombre de *flvC*, comienza en la posición 31574 con un codón de inicio ATG y un probable sitio de unión al ribosoma GAGGAGAAGG situado 26 nucleótidos corriente arriba del inicio (Anexo I). Se trata de un gen de aproximadamente 19 kb que no está secuenciado en su totalidad. Presenta un fragmento de ADN interno sin secuenciar de 3200 pb que hace que la secuencia estudiada en este trabajo se separe en dos anexos. Este gen se estima que termine en la posición 12522 con un codón de terminación TTG (Anexo II).

ORF6: Se trata de una región abierta de lectura incompleta dónde sólo está presenta la parte codificante para el extremo N-terminal de la proteína que de dicho gen se deduce. Comienza en la posición 12555 con un codón de inicio ATG precedido de un probable sitio de unión al ribosoma GAAAGG situado 9 nucleótidos corriente arriba (Anexo II). Se lo denomina *flvD*.

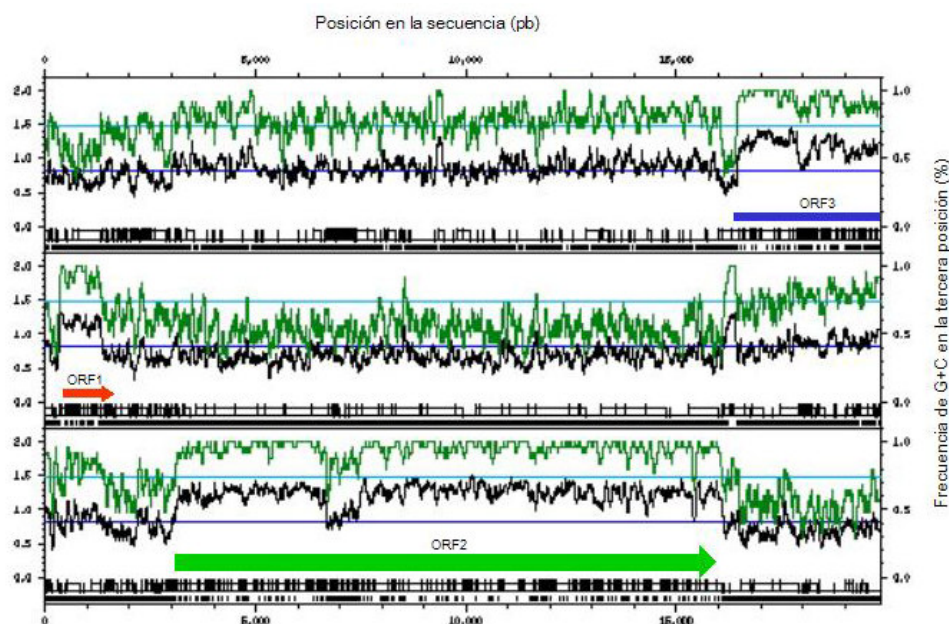


Fig. 36. Gráfico generado por el programa *CODONPREFERENCE* (GCG), donde se muestran las regiones abiertas de lectura presentes en parte de la secuencia contenida en el Anexo I. Dicho programa utiliza la tabla de preferencias en el uso de codones de *Streptomyces* creada por Wright y Bibb (1992). Con flechas de colores se han marcado las regiones abiertas de lectura ORF1, ORF2 y ORF3.

3.5. FUNCIONES DEDUCIDAS DE LAS REGIONES ABIERTAS DE LECTURA DE LA PKS DE FLAVOFUNGINA I

La comparación de la secuencia de aminoácidos deducida de cada una de las fases abiertas de lectura con proteínas de la base de datos, permitió asignar una posible función para cada una de ellas que se presenta a continuación en la tabla 13:

ORF	Polipéptido	Posición en la secuencia	RBS	Función propuesta
ORF1	Orf1	⁹³⁸ ATG/TGA ³⁶³	GAAAAGAGG	Reguladora
ORF2	Orf2	¹¹¹⁹ ATG/TAG ²⁰⁸⁷	AAGGAGAA	Esterasa
ORF3 (<i>flvA</i>)	FlvA	³⁸⁷² ATG/TGA ¹⁶⁹⁸⁷	AGAGG	PKS Tipo I
ORF4 (<i>flvB</i>)	FlvB	¹⁷²⁰⁴ TGT/TGA ³¹⁴¹²	GAGGA	PKS Tipo I
ORF5 (incompleta) (<i>flvC</i>)	FlvC	³¹⁵⁷⁴ ATG/TTG ¹²⁵²²	GAGGAGAAGG	PKS Tipo I
ORF6 (incompleta) (<i>flvD</i>)	FlvD	¹²⁵⁵⁵ ATG	GAAAGG	PKS Tipo I

Tabla 13. Funciones deducidas de las ORFs presentes en la secuencia estudiada.

ORF1: región abierta de lectura que codifica una proteína de 192 aa con un peso molecular estimado de 19,5 kDa. Presenta una alta homología con proteínas reguladoras de la transcripción y más en concreto con aquellas pertenecientes a la familia TetR. El porcentaje más alto de homología lo exhibe con un regulador transcripcional presente en *Nocardia farcinica* (EMBL YP_119011). También presenta homología con un regulador transcripcional de la familia TetR que se encuentra en *Streptomyces avermitilis* (EMBL BAC70165), entre otros.

```

Score = 218 bits (554), Expect = 3e-55
Identities = 103/138 (74%), Positives = 118/138 (85%), Gaps = 0/138 (0%)
Frame = -2

Orf1      KSTVYRRWPSK GELIVDCLLD AFGPAPLTGSSRAELMTSTIRWVAAKIGEPGVGDAFAGV
K+TVYRRWPSKG LI+DCLLD AFGP PL G+ RAE+MT IRW+A KI EPGVGD AFAAGV
R.T.      KTTVYRRWPSKG ALIIDCLLD AFGPVPLEGADRAEIMTGAI RWIAGKIAEPGVGDAFAGV

Orf1      FSDAVSDPALREILSTR LQDPYRLALQEALDEPENRVLFFIDVVVGVLLHRMGMTGPEMV
FSDAVSDP LR +LSTR QDPYR+A+Q+AL E E RVL F+IDVVVG LLHRMGMTG PMV
R.T.      FSDAVSDPELR TLLSTRFQDPYRIAVQDALGESEQRVLFYIDVVVGTLLHRMGMTGAPMV

Orf1      EADV DALVDMVLP HFTDG
E+DV+AL++M L +F G
R.T.      ESDVEALIE MTLRYFGTG

```

Fig. 37. Alineamiento de la proteína Orf1 con la proteína presente en la regulación de la transcripción de *Nocardia farcinica* (R.T.).

ORF2: región abierta de lectura que codifica una proteína de 321 aa cuyo peso molecular es de 34,3 kDa homóloga a proteínas de tipo esterasas. La identidad máxima, un 36%, la presenta con la esterasa de *Ralstonia solanacearum* (EMBL NP_521793). Concretamente, se ha encontrado homología de esta esterasa con tioesterasas Tipo II. En la figura 38 se muestra el alineamiento de la proteína Orf2 con proteínas tioesterasas Tipo II, donde se señala el motivo conservado GX SXG alrededor del residuo serina, así como el residuo conservado histidina (H) aproximadamente 100 aminoácidos corriente abajo del centro activo.

```

                GX SXG                                H
Sac      DGPFALFGHSMGAL IAYETARRLAA-111aa-TRTFPGGHFYL
Sfr      GVPVALFGHSMGAVVAYETARRLAA-113aa-LRVFPGGHFYL
ScoT     DRPHALFGHSMGAL IAYELARLAA-113aa-LRVLPGGHFYL
Amy      DRPLALFGHSMGAL IGYELALRMAA-113aa-LRVLPGGHFYL
Orf2     DPERTAVVEGFSGAL I SAI AATRAA-121aa-TEHAEAPHAATL

```

Fig. 38. Alineamiento de la secuencia de aminoácidos de la proteína Orf2 de *S. sp.* CO-AC25 con la tioesterasas Tipo II: Sac (*Saccharopolyspora erythraea* (EMBL X60379)), Sfr (*Streptomyces fradiae* (EMBL U08223)), ScoT (*Streptomyces coelicolor* (EMBL AF109727)) y Amy (*Amycolaptosis mediterranei* (EMBL AF040571)).

ORF3: a la proteína deducida a partir de esta región abierta de lectura se la ha denominado FlvA, tiene un tamaño de 4329 aa y un peso molecular de 472,8 kDa. La comparación de la secuencia de aminoácidos de FlvA con la de otras proteínas de la base de datos muestra homología con proteínas PKSs Tipo I como la de *Streptomyces aizunensis* (EMBL AAX98191) o la de *Streptomyces noursei* (EMBL AAF71767), con las que presenta una

identidad del 53%, o la de *Streptomyces nodosus* (EMBL AAK73514), con la que tiene una identidad del 50%.

De dicha comparación se puede deducir la presencia de dominios catalíticos KS, AT, KR, DH y ACP. El extremo N-terminal de esta proteína presenta los dominios funcionales AT y ACP, arquitectura que sugiere que pueda tratarse del módulo de “carga” de la ruta biosintética que en este caso iría seguido de dos módulos de elongación.

El dominio AT presente en el posible módulo de “carga” exhibe una identidad máxima del 47% con el dominio AT de la proteína NysB de *Streptomyces noursei* (EMBL AAF71775). Este dominio AT presenta el motivo conservado GSSAG alrededor del residuo serina del centro activo (Anexo I), a diferencia de los dominios AT presentes en las PKS de arranque de otros macrólidos polienos caracterizados, como nistatina (Brautaset *et al.*, 2000), pimaricina (Aparicio *et al.*, 1999; Aparicio *et al.*, 2000), anfotericina B (Caffrey *et al.*, 2001) o candicidina (Campelo y Gil, 2002), donde el motivo conservado es GHSXG (Figura 39).

PimS2	VVGH	SGGE	I	AAAA	V	V	A	G	A	L	S	L	E	D	A	A	R	V	V	A	L	R	S	Q	A	I	A	R	G	L	A	G	T	G	G	M	L	S	V	P	L	P	A	A	D	V	E	Q	R	L	A									
FscD	V	I	G	H	S	Q	G	E	I	A	A	A	V	S	G	A	L	I	L	D	G	A	R	V	V	A	L	R	S	Q	A	I	R	R	V	L	A	G	A	G	G	M	S	V	Q	V	S	A	A	F	A	E	Q	Y	L	T	P			
EryAIII	V	V	G	H	S	Q	G	E	I	A	A	A	H	V	A	G	A	L	I	L	E	D	A	A	K	L	V	A	V	R	S	R	V	L	R	R	-	L	G	G	Q	G	M	A	S	F	G	L	G	T	E	Q	A	A	E	R	I	G	R	
NysB	V	V	G	H	S	Q	G	E	I	A	A	A	C	V	A	G	A	L	S	L	A	D	A	A	R	V	V	A	L	R	S	Q	A	L	P	Q	-	L	S	G	R	G	G	M	S	V	S	A	P	V	E	R	V	T	A	L	L	A	P	
AmphB	V	V	G	H	S	Q	G	E	I	A	A	A	H	V	A	G	A	L	S	L	E	D	A	A	L	V	V	A	L	R	S	Q	A	L	P	Q	-	L	S	G	L	G	G	M	S	V	S	A	P	A	A	R	V	T	E	L	L	O	P	
FlvA	V	I	G	S	S	A	G	E	L	P	A	A	T	V	A	G	A	L	S	V	E	D	G	A	R	A	V	A	V	F	G	Q	A	Q	V	H	-	L	L	E	H	G	V	M	L	S	V	L	A	S	A	D	E	V	R	E	R	L	T	R

Fig. 39. Alineamiento de los dominios AT de FlvA y otras proteínas de la base de datos que intervienen en las rutas de biosíntesis de diferentes antibióticos como pimaricina (PimS2), FR-008 (FscD), eritromicina (EryAIII), nistatina (NysB) y anfotericina (AmphB).

La secuencia de aminoácidos que antecede a dicho centro activo, RPDVGQPALWAMMVSLAALW-11aa-GSS, sugiere que la unidad iniciadora en el proceso de biosíntesis del macrólido oxopentaeno flavofungina I sea un derivado del propionato o metilmalonil-CoA. En base a la estructura química del oxopentaeno se postula que dicha unidad iniciadora sea metilpropionil-CoA o isobutiril-CoA.

El dominio ACP del posible módulo de “carga” de la proteína FlvA presenta un grado de identidad máximo del 61% con la proteína PimS2 de *Streptomyces natalensis* (EMBL CAC20921), y en él se observa el motivo conservado característico de los dominios ACP (LGVDSA) (Anexo I).

Los dos posibles módulos de elongación que siguen al hipotético módulo de “carga” se organizan según se muestra en la figura 40. Ambos presentan un porcentaje de identidad con las proteínas NysJ de *Streptomyces noursei* (EMBL AAF71767) y AmphC de *Streptomyces nodosus* (EMBL AAK73514) del 53% y 52% respectivamente.

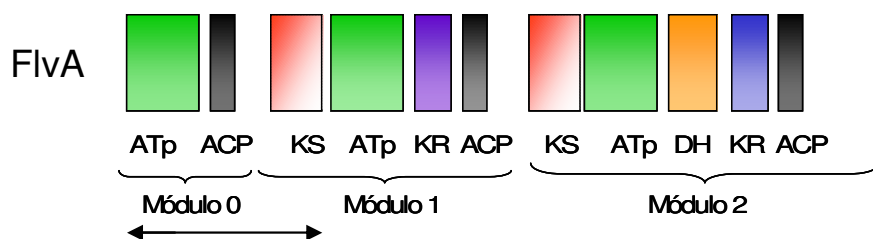


Fig. 40. Estructura modular de la proteína deducida FlvA, en la que se muestra el módulo de “carga” (módulo 0) seguido de dos módulos de elongación (módulos 1 y 2). Con flechas se señala el fragmento empleado para realizar la interrupción génica (pLM31).

Todos los dominios catalíticos identificados muestran los motivos conservados característicos (Anexo I).

Así, los dominios catalíticos KS presentan el motivo conservado GPAVTVDTACSSS, alrededor del residuo cisteína y contienen los residuos invariables histidina (H) 135 y 175 aminoácidos en dirección C-terminal de dicho residuo cisteína, como sucede con los dominios KS descritos para la PKS de pimarcina (Aparicio *et al.*, 1996; Aparicio *et al.*, 1999).

Los dominios AT de ambos módulos de elongación presentan el motivo conservado GHSXG alrededor del residuo serina y en los dos casos la secuencia de aminoácidos que determina la unidad de elongación incorporada se corresponde con la que incorpora metilmalonil-CoA o propionato (ATp).

Los módulos de elongación de la proteína FlvA se diferencian en los dominios catalíticos implicados en el grado de reducción de la cadena policetónica. Mientras que para el módulo de elongación 1 sólo aparece el dominio catalítico KR, para el módulo de elongación 2 también se observa la presencia de un dominio DH.

El dominio KR del primer módulo de elongación presenta el motivo conservado GTGAIGGHVA y el dominio KR del segundo módulo presenta el motivo GLGTLGAIVA, siendo la secuencia consenso para estos dominios funcionales GXGXXGXXXA.

El dominio funcional DH presenta el motivo conservado característico HGVGGRVLPF.

	HXXXGXXXXP
FlvA2	FTGRLSTRTHPWLAD HGVGGRVLPF PGTGFLDLALHAASHCGLDVTDELTLQAPMALPERG
AmphC	FTGRLSLQSHPWLAD HAVHGTVLLP GTALLDLAFRAGDEVGCDRVEELTLAAPLVLPERG
NysC	-----SVRTHPWLAD HGVQGRALLP GTAFVELAVRAGDEAGCDRVEELTLAAPLVLPERG
PimS3	LTGRLSLQTHPWLT DAIGGNVFFP GTGFLELAMHAADQAGCDRVEDLVLMTPLVLAEN

Fig. 41. Alineamiento de la secuencia consenso del dominio DH de la proteína deducida FlvA junto con otros dominios DH de las proteínas AmphC de anfotericina (*S. nodosus*), NysC de nistatina (*S. noursei*) y PimS3 de pimarcina (*S. natalensis*).

Algunos de los motivos conservados presentes en la proteína FlvA se encuentran reflejados en la figura 42.

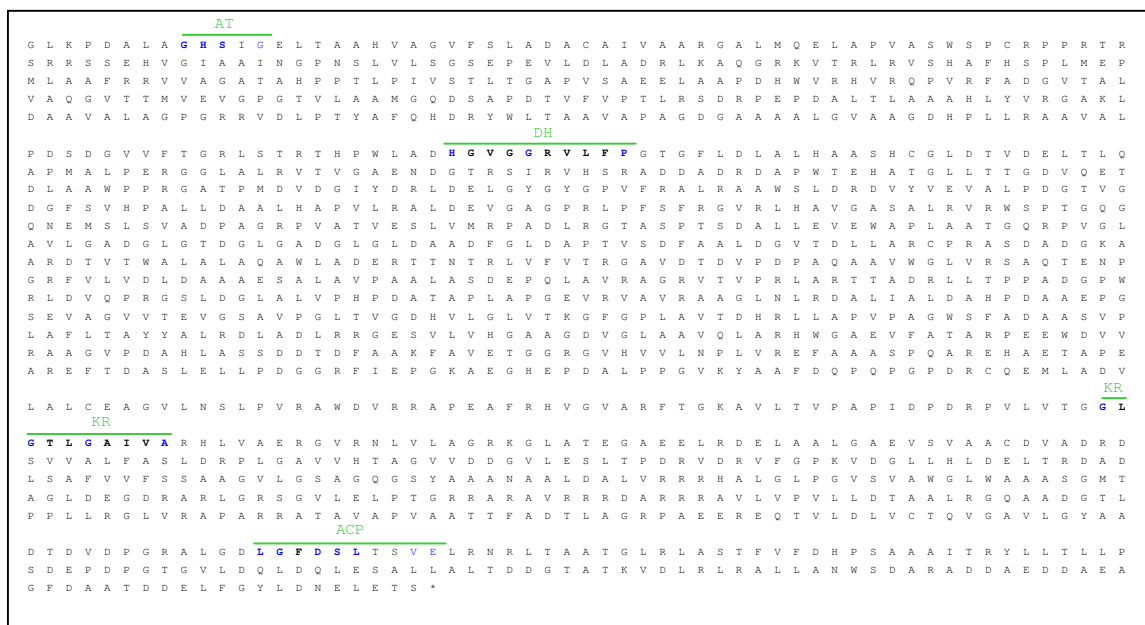


Fig. 42. Secuencia de aminoácidos correspondiente a parte del segundo módulo de elongación de la proteína FlvA. En azul se marcan los residuos conservados para cada dominio.

ORF4: codifica una proteína de 4736 aa a la que se le ha dado el nombre de FlvB y cuyo peso molecular es de 479 kDa aproximadamente. La comparación de su secuencia de aminoácidos con otras proteínas de las bases de datos manifiesta su alta homología con PKSs Tipo I. Así, aparece un grado de identidad del 56% con la PKS Tipo I de *Streptomyces aizunensis* (EMBL AAX98186) y del 52% con la proteína PimS2 de *Streptomyces natalensis* (EMBL CAC20921) o con la proteína Amphi de *Streptomyces nodosus* (EMBL AAK73501).

ORF5: codifica una proteína denominada FlvC. La comparación de esta proteína con proteínas de las bases de datos revela un grado de identidad del 57% con la PKS Tipo I de *Streptomyces aizunensis* (EMBL AAX98184 y AAX98189).

ORF6: codifica una proteína a la que se ha denominado FlvD. Dicha proteína está incompleta y se ha secuenciado únicamente en su extremo N-terminal, localizando en dicho extremo un módulo de elongación de la cadena policetónica de flavofungina I en el que solamente se puede distinguir el dominio catalítico KS.

La proteína FlvB esta constituida por 3 módulos de elongación secuenciados en su totalidad, mientras que la proteína FlvC se compone de al menos 4 módulos de condensación, de los cuales dos están secuenciados por completo. De todos los módulos de condensación que constituyen ambas proteínas se han podido distinguir sus dominios catalíticos, colocados en el orden preciso en el que van a ser utilizados durante la síntesis de la cadena policetónica y cuya disposición se muestra en la figura 43.

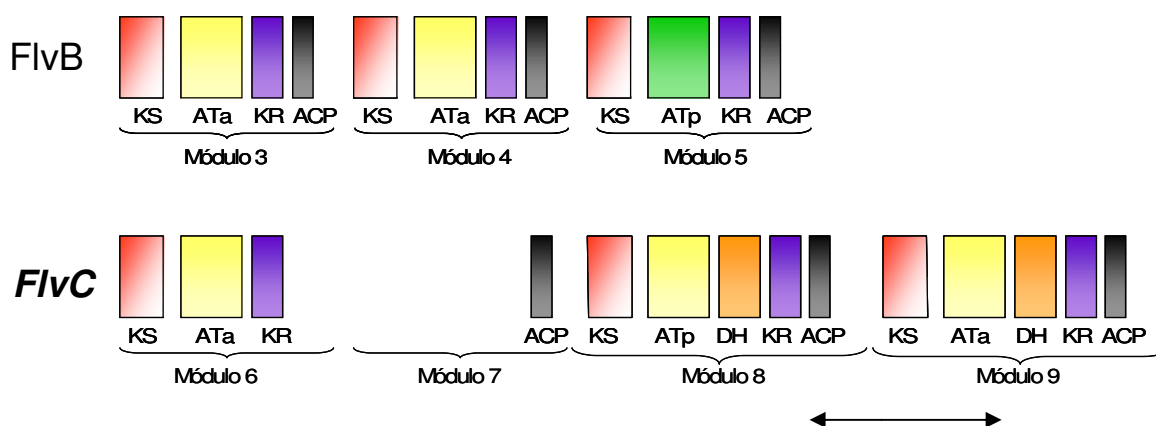


Fig. 43. Estructuras modulares de las proteínas deducidas FlvB y FlvC. Cada módulo consta de un número variable de dominios catalíticos. Con flechas se señala el fragmento empleado para realizar la disrupción génica (pLM32).

Debido al alto grado de conservación que presentan las secuencias de aminoácidos de estos dominios se ha procedido a un análisis conjunto de los mismos.

Dominios KS: presentan la secuencia altamente conservada GPAVTVD**TACS**SSL y contienen el residuo invariable histidina alejado 135 y 175 aminoácidos en dirección C-terminal del residuo cisteína, al igual que ocurría con los dominios KS de la proteína FlvA y con los de los dominios KS de Pim03 y Pim04 de pimarcina (Aparicio *et al.*, 1999).

	GPXXXX TACS S			
FlvB3	NGSAGSIASGRIAYT	LGLV GPAVTVDTACS	SSLVALHLMVALRR	EECSLALAGGATFMA
FlvB4	TGSAASVISGFGRSL	YTLE GPAVTVDTACS	SSLVALHLAVQALQK	GECELALAGGVTVM
FlvC6	TGISASVISGRLSYT	FGLE GPAVTVDTACS	SSLVALHLAAQALHK	GECDMALVGGVTVM
FlvC8	TGSAASVISGRLSYT	FGLE GPAVTVDTACS	SSLVALHLAGQALNK	GECELALVGGVTVM
FlvC9	TGNAASVVSGRVAYA	FGLE GPAVTVDTACS	SSLVALHMAVQALRT	GDCSMALAGGVTVM
FlvD	TGNAASVVSGRISYV	LGLE GPALTIDTACS	SSLVALHLAVHALRH	RECGLALAGGVTVM
FscE	TGLAASVVSGRLAYT	FGLE GPAVTVDTACS	SSLVALHWAMQALRA	GECDLAVAGGITMMS
NysJ	TGLAASVISGRLAYA	FGFE GPAVTVDTACS	SSLVALHWAVQALRA	GECSLALAGGVTVM

Fig. 44. Alineamiento de las secuencias conservadas de los dominios catalíticos KS de las proteínas deducidas a partir de la ORF4 (FlvB), ORF5 (FlvC) y ORF6 (FlvD) junto con los dominios KS de las proteínas FscE de *S. sp.* FR-008 y NysJ de *S. noursei*. En azul se señala el residuo cisteína del centro activo.

Dominios AT: presentan todos el motivo conservado GH**SXG** alrededor del residuo serina. Dichos dominios se pueden englobar en dos grupos, por un lado los que incorporan metilmalonil-CoA (propionato) (Marsden *et al.*, 1994) y los que incorporan malonil-CoA (acetato) (Haydock *et al.*, 1995).

Así los dominios AT que incorporan metilmalonil-CoA (propionato) (ATp) presentan la secuencia consenso RVDVV-----M-S-A—W-11aa-GH**SXG**, mientras que aquellos dominios AT que incorporan malonil-CoA (acetato) (ATa) presentan la secuencia consenso ETGYA-----Q-A-FGLL-11aa-GH**SXG**.

GHSXG

```

FlvC6  MERLDVVQPALFAVM  VSLAATWQSWGVEPS  AVLGHSQGEIAAACV  AGALPLKDAARLVVL
FlvD   LERVDVVQPALFSVM  VSLAALWRSYGVEPD  AVVGHSQGEIAAAVY  AGALSLADAACKIVAQ
NysB   LERVDVVQPALFAMM  VGLSALWRSYHGVPV  AVVGHSQGEIAAACV  AGALSLADAARVVAL
FscB   LERVDVVQPVLWAVM  VSLAEVWRAQGVPEP  AVIGHSQGEIAAAVV  AGGLSLEDGARVVAL
FlvB3  LDRTEYAQPALFAVE  VALYRLWESWGKPD   ALAGHSIGELTAAHV  AGVFSLADACAIVAA
FlvB4  LNETGVTQPALFAVE  VALFRLLSWGVRPD   VLAGHSIGELAAAHV  AGVWSLADAACKVVAA
FlvB5  LNETGVTQPALFAVE  VALFRLLSWGVRPD   VLAGHSIGELAAAHV  AGVWSLADAACKVVAA
FlvC8  LNETGVTQPALFAVE  VALFRLLSWGVRPD   MLAGHSIGELAAAHV  SGVWSLPDAACKIVAA
FlvC9  LNETGVTQPALFAVE  VALFRLLSWGVRPD   MLAGHSIGELAAAHV  SGVWSLPDAVKVSA

```

Fig. 45. Alineamiento de las secuencias altamente conservadas de los dominios AT presentes en las ORF4 (FlvB), ORF5 (FlvC) y ORF6 (FlvD) de *S. sp. CO-AC25* junto con las AT de nistatina (NysB) de *S. noursei* y de FR-008 (FscB) de *S. sp. FR-008*. En azul se señala el residuo serina del centro activo y en verde la secuencia consenso de aminoácidos que determina la incorporación de acetato o propionato por parte de la AT.

Dominios KR: en todos los casos aparece el motivo conservado de unión al NADP(H) GXGXXGXXXA, salvo para el dominio KR del módulo 8 de la proteína FlvC cuyo motivo conservado es GXXGXXGXXXA.

GXGXXGXXXA

```

FlvB5  TADSTVLITGGLGVLGGLTARHLVTRHGVG
FlvB3  SPRGTVLVTGGTGALGGHVARWVAAH-GAE
FlvB4  TPRGTVLVTGGTGALGGHVARWLAA-----
FlvC6  SPRGTVLVTGGTGALGAHVARWLVAK-GAE
pim03  ---GTVLITGGTGGLGRSVARHLVSEHGVR

```

Fig. 46. Alineamiento de las secuencias consenso características de los dominios KR presentes en las proteínas deducidas FlvB y FlvC de *S. sp. CO-AC25* y la proteína Pim03 de pimarcina (*S. natalensis*).

Dominios DH: la comparación de las secuencias de aminoácidos presentes en la proteína FlvC inmediatamente anteriores a los dominios KR con la de otras proteínas de las bases de datos, revelaron la existencia de dominios DH en los cuales no se ha localizado la existencia de motivo conservado alguno. Estos dominios catalíticos podrían ser necesarios teniendo en cuenta la estructura química de la flavofungina I y su mecanismo biosintético.

Dominios ACP: en todos ellos se encuentra el motivo conservado LGXD**SL**XXVE, alrededor del residuo serina.

LGXDSL**XXVE**

```

FlvB3  RQLTELVRYEAAAVL  GHASGDAFSATRAFR  DLGFDSLTAVELRAR  LSEATGVALPAAVVF
FlvB4  RQLTELVRYEAAAVL  GHVSGDAFSATRAFR  DLGFDSLTAVELRAR  LSDATGLALPATLVF
FlvB5  RTLVLVRSHTAAVL  GFAGPEAVEAERAFK  ELGFDSLTGVELRNR  LDAEVGTRLPATLVF
FlvC7  KALLDLVLGHVAVVL  GHSSAGAIAPDRAFK  DLGFDSLSSVELRNH  LGAATELALPATLVF
FlvC8  KALLDLVLEHVAVVL  GHGSAGAIVPDRAFK  ELGFDSLTAVELRNR  LGAATELRLSATLVF
FlvC9  RVVLDLVREQVAVL  GYDSSTQVEPGRAFK  DLGFDSLTAVEFRNV  FNSATGLSLPSTLVF
FscC   EVLLDLVRGQIAVVL  GHSGAQTVNPHRAFQ  DLGFDSLTAVELRNR  LGKVTGLRLPATVF
AmphC  DALLTLVREQAALVL  GHSGGGGIDPSRAFR  DLGFDSLTAVELRNR  LGAATGVRLSATVF

```

Fig. 47. Alineamiento de las secuencias consenso de los dominios ACP de las proteínas FlvB, FlvC de *S. sp. CO-AC25* y los dominios ACP de las proteínas FscC de *S. sp. FR-008* y AmphC de *S. nodosus*. En azul se marca el residuo serina.

3.6. ANÁLISIS TRANSCRIPCIONAL DE LOS GENES *flvA* Y *flvB*

Con objeto de iniciar el estudio de la transcripción de los genes clonados del *cluster* de flavofungina I, se procedió a realizar ensayos de protección frente a endonucleasa S1 (Materiales y Métodos, apartado 2.15).

Análisis transcripcional del gen *flvA*

Para llevar a cabo dicho estudio, se extrajo el ARN de la cepa productora a diferentes tiempos (24 h, 30 h, 48 h, 54 h y 96 h) y con ellos se realizaron los ensayos de protección frente a endonucleasa S1 de alta resolución.

La sonda utilizada se generó mediante una reacción de amplificación con los oligos diseñados FlvA-I y FlvA-II y el plásmido pF1 como molde. El oligonucleótido FlvA-I tiene una longitud de 21 pb y se encuentra entre las posiciones 3624 y 3644 (secuencia Anexo I) y el oligonucleótido FlvA-II también tiene una longitud de 21 pb y se sitúa entre las posiciones 3950-3930 (secuencia Anexo I), éste último se marcó con [32 P]dATP en su extremo 5'. El plásmido pF1 se trata de un fragmento *Bam*HI del cromosoma del productor clonado en el pIJ2925 (nucleótidos 1815-3957 del Anexo I) (Figura 48).

El fragmento de PCR obtenido tiene un tamaño de 327 pb y solapa con el extremo 5' del gen *flvA*.

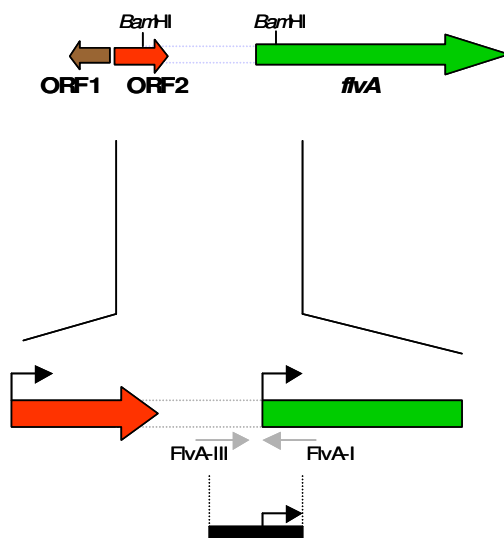


Fig. 48. Esquema de la sonda generada al amplificar con los oligos FlvA-I y FlvA-II sobre el plásmido pF1. En negro se representa la sonda utilizada para el ensayo de protección con endonucleasa S1.

El ensayo de S1 mostró un fragmento protegido de 175 pb, 152 pb menor que la sonda, lo que permitió identificar un probable inicio de transcripción en la posición 3776 (A) o en la posición 3777(G), a 94 ó 93 pb del inicio de traducción del gen *flvA* (ATG).

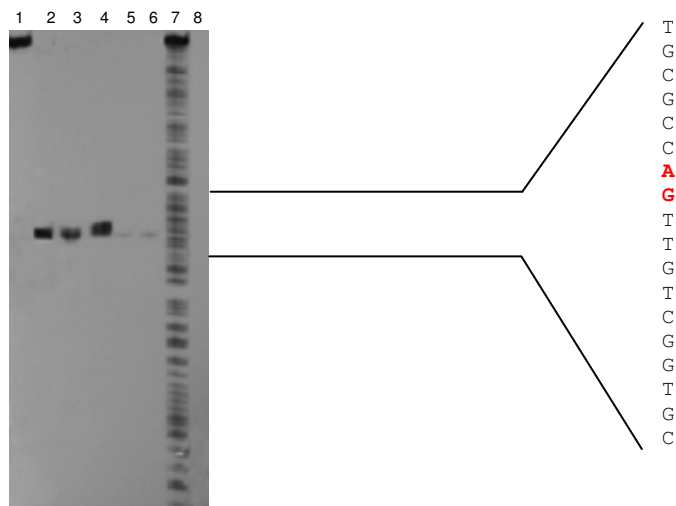


Fig. 49. Determinación del origen de transcripción de *flvA* (1) Sonda; (2)→(6) Fragmentos protegidos obtenidos utilizando los ARNs de 24 h (2), 30 h (3), 48 h (4), 54 h (5) y 96 h (6). (7) Reacción de secuencia de Maxam y Gilbert (1980) (8) Control: sonda tratada con S1.

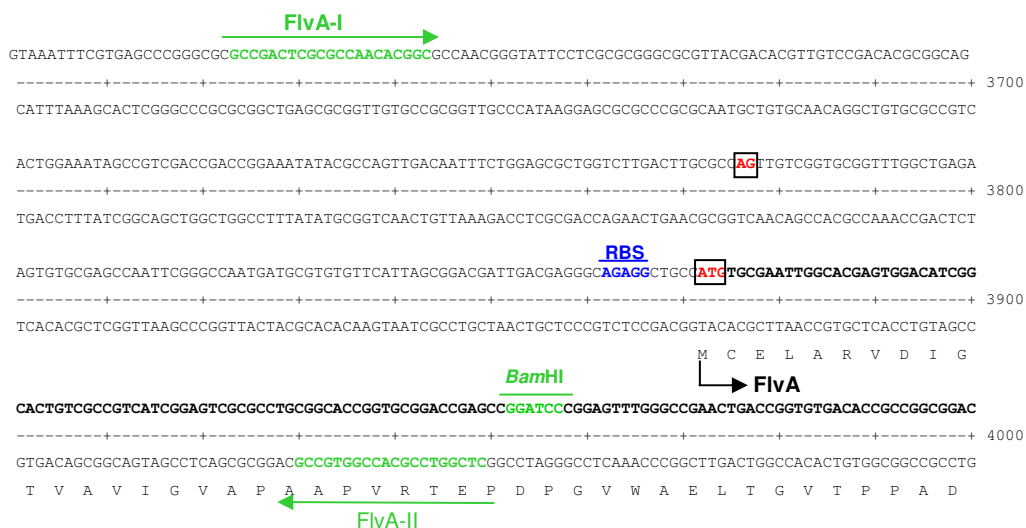


Fig. 50. Secuencia de nucleótidos solapante con el extremo 5´ del gen *flvA*. Se muestran los oligonucleótidos FlvA-I y FlvA-II; el probable RBS del gen *flvA*; y el inicio de traducción (ATG) y transcripción (A ó G) de dicho gen.

En paralelo al análisis transcripcional del gen *flvA*, se realizó una curva de producción de flavofungina I en medio sólido R5, con el fin de poder correlacionar los niveles de expresión del gen con los niveles de producción del macrólido oxopentaeno (Figura 51). Se aprecia un desfase entre la expresión y la producción de metabolitos.

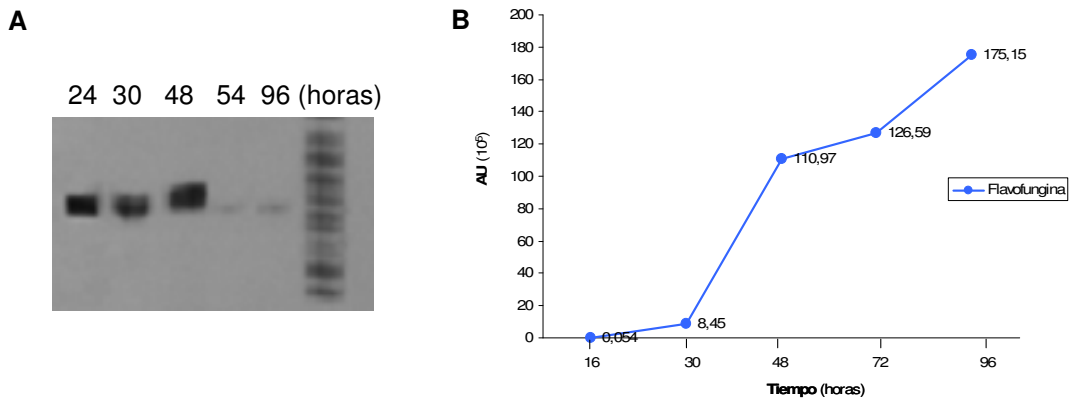


Fig. 51. A. Expresión temporal del gen *flvA*. **B.** Curva de producción del macrólido oxopentaeno flavofungina I producido por *Streptomyces* sp. CO-AC25 en medio sólido R5.

Análisis transcripcional del gen *flvB*

El gen *flvB* está codificado corriente abajo del gen *flvA*. Con el fin de determinar si se transcribe como un policistrón, se realizó un ensayo de protección a endonucleasa S1 de baja resolución.

Como sonda para este ensayo se utilizó un fragmento amplificado de PCR a partir de los oligonucleótidos diseñados FlvB-I y FlvB-II y el plásmido pG41 como ADN molde (fragmento *Pst*I del cromosoma en el plásmido pIJ2925). El fragmento amplificado se clonó en el sitio de restricción *Hin*clI del vector pUC18 y a la construcción resultante se la denominó pUC-FlvB. Para usarlo como sonda en el ensayo de protección con endonucleasa S1 se rescató como *Eco*RI/*Nde*I. Dicha sonda posee el extremo 3' terminal del gen *flvA* (32 pb), la región intergénica *flvA-flvB* (215 pb), el extremo 5' del gen *flvB* (84 pb) y un fragmento de ADN heterólogo, fragmento *Nde*I-*Hin*clI (213 pb) del plásmido pUC18, que servirá de control de la digestión de la endonucleasa S1 (Figura 52).

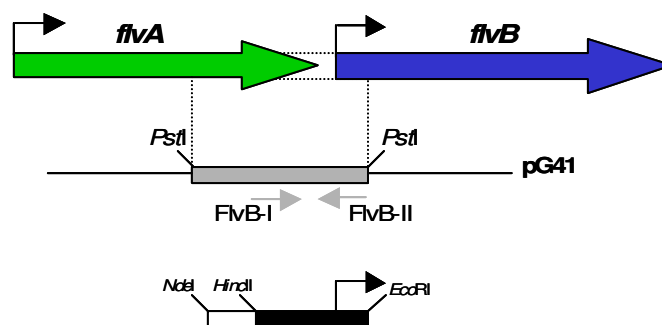


Fig. 52. Esquema de la sonda generada al amplificar con los oligonucleótidos FlvB-I y FlvB-II sobre el plásmido pG41 y posterior clonaje en el pUC18. En color negro se representa el fragmento amplificado y en color blanco el fragmento heterólogo *Nde*I/*Hin*clI.

La sonda se marcó con [α^{32} P]dGTP (Materiales y Métodos, apartado 2.12) y permitió identificar un fragmento protegido de tamaño similar al amplificado por PCR pero inferior a la sonda utilizada. Ello sugiere que el gen *flvB* forma parte de un policistrón junto a *flvA* (Figura 53).

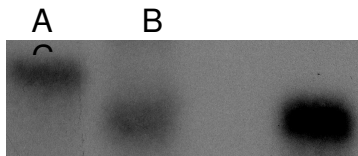


Fig. 53. Análisis transcripcional del gen *flvB*. **A.** Sonda empleada en el ensayo de protección frente a endonucleasa S1. **B.** Fragmento de PCR amplificado. **C.** Fragmento protegido, de tamaño similar al fragmento de PCR amplificado.

3.7. GENERACIÓN DE UN MÓDULO DE “CARGA” HÍBRIDO EN LA PROTEÍNA FIV A

Con el objeto de producir nuevas moléculas recombinantes se llevó a cabo una sustitución del dominio catalítico aciltransferasa (AT) del módulo de “carga” de FivA por un dominio AT heterólogo.

Para esta sustitución se eligió el dominio AT del módulo de “carga” de la PKS implicada en la biosíntesis de los macrólidos tetraenos rimocidina y CE-108, ambos producidos por *Streptomyces diastaticus* var. 108 (Seco *et al.*, 2004) (RimA). El módulo AT de RimA es flexible en el reconocimiento de sustrato, por lo que la proteína RimA es capaz de incorporar indistintamente acetil-CoA o butiril-CoA originando bien CE-108 o bien rimocidina. Los cambios esperados en la estructura química de flavofungina I (Figura 54.A) como resultado de dicha sustitución tendrían lugar a nivel de la cadena lateral del C-31, convirtiéndose ésta en una cadena lateral metilo (Figura 54.B) o propilo (Figura 54.C) en caso de ser incorporadas acetyl-CoA o butiril-CoA respectivamente como unidades iniciadoras.

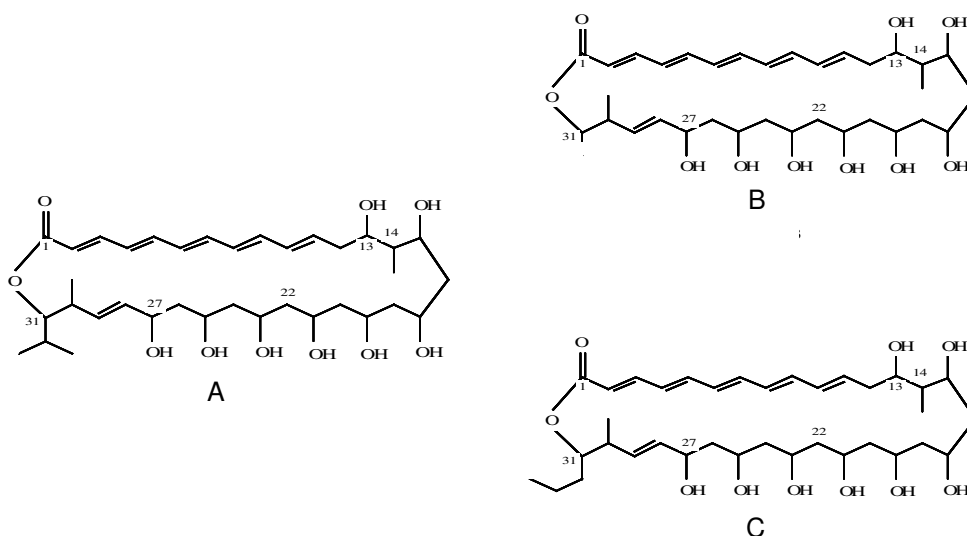


Fig.54. **A.** Estructura química de flavofungina I. **B, C.** Estructuras químicas resultantes de la sustitución del dominio AT nativo por el de la proteína RimA.

La sustitución del dominio catalítico AT nativo por el dominio catalítico AT heterólogo se realizó utilizando la conjugación como sistema para generar “*in vivo*” una PKS híbrida que previamente ha sido modificada “*in vitro*”. Para ello se realizó una construcción sobre un plásmido conjugativo y suicida al que se denominó pME, donde el dominio catalítico AT del módulo de “carga” de la proteína FlvA estaba sustituido por el del *cluster* de biosíntesis de rimocidina y CE-108, y que por doble recombinación, se integró en el cromosoma de *S. sp.* CO-AC25 originando los cambios esperados (Figura 55).

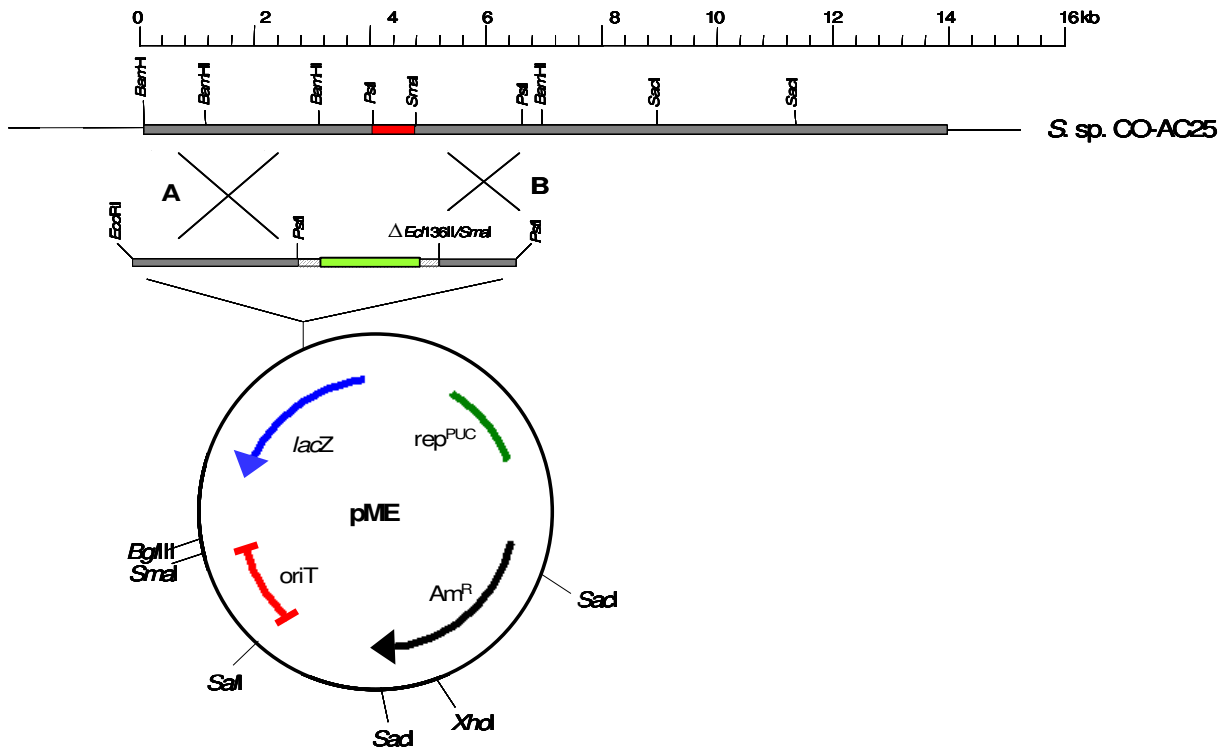


Fig. 55. Esquema de la recombinación homóloga ocurrida entre el plásmido pME y el cromosoma de *S. sp.* CO-AC25.

Construcción del gen *flvA* híbrido (*flvA**)

Para la obtención de dicho gen fue necesario realizar varias construcciones:

- **Generación del plásmido p11J:** plásmido resultante del clonaje del fragmento *EcoRI* del fago recombinante EMBL4/11J de aproximadamente 14 kb proveniente en el vector pIJ2925. Dicho fragmento contiene la secuencia codificante del gen *orf1*, que codifica una esterasa y la casi totalidad del gen *flvA* (nucleótidos 794-14940, Anexo I) (Figura 56).

Como resultado de la amplificación se obtuvo un fragmento de aproximadamente 750 pb portador del dominio catalítico AT correspondiente al RimA flanqueado por los sitios de restricción *Pst*I (corriente arriba) y *Sac*I (corriente abajo) generados en la amplificación y que facilitarían posteriores clonajes a la hora de generar “*in vitro*” la PKS híbrida. El fragmento amplificado fue romado mediante la actividad enzimática del fragmento Klenow de la ADN polimerasa y, a continuación, se clonó en el vector pUC18 en el sitio de restricción *Hinc*II; así se obtuvo el plásmido pAT (Figura 58).

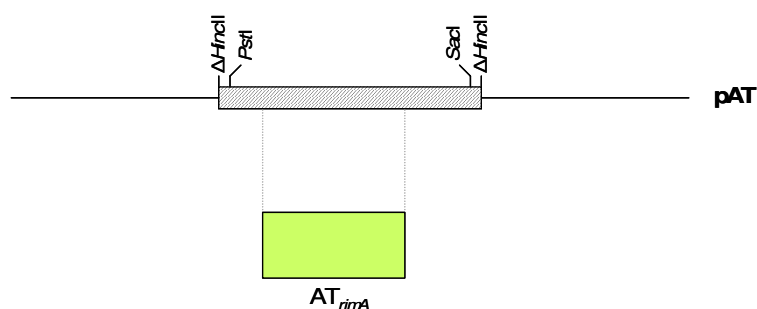


Fig. 58. Esquema de la construcción del plásmido pAT. Se señalan los sitios de restricción diseñados *Pst*I y *Sac*I.

- **Generación del plásmido pASAT:** a partir de los clones p11J y pAT se obtuvo el plásmido pASAT. El plásmido p11J fue digerido con las enzimas *Bam*HI-*Sma*I proporcionando un fragmento de 2,5 kb que contiene el dominio catalítico ACP del módulo de “carga” de FlvA y los dominios KS y AT del primer módulo de elongación de la PKS de flavofungina I (nucleótidos 8154-5563, Anexo I). A su vez, el plásmido pAT fue digerido con las enzimas de restricción *Ec*I136II (que reconoce la secuencia *Sac*I liberando extremos romos) y *Pst*I, liberando el fragmento amplificado de PCR que contiene el dominio AT del módulo de arranque RimA. Ambos fragmentos se clonaron en el vector pUC18 en los sitios de restricción *Bam*HI-*Pst*I originando el plásmido pASAT (Figura 59).

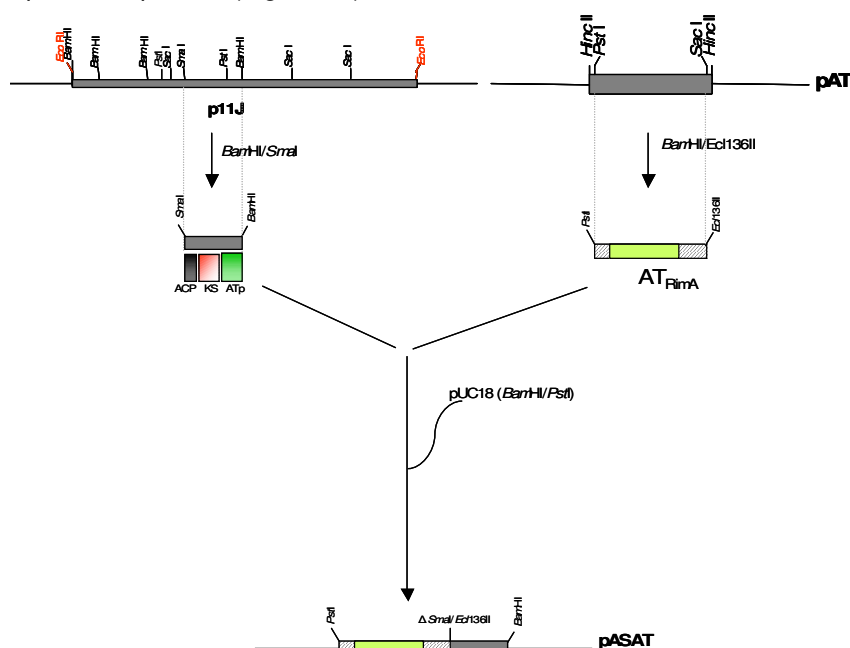


Fig. 59. Esquema de la construcción del plásmido pASAT a partir de las construcciones p11J y pAT.

- **Obtención del plásmido pME:** finalmente, una vez obtenido el plásmido pASAT se digirió con *Pst*I y junto con el fragmento *Eco*RI-*Pst*I de aproximadamente 4 kb del clon p11J que codifica los genes *orf1* y el extremo 5' terminal del gen *flvA* (nucleótidos 794-4754, Anexo I) ambos fragmentos se clonaron en el vector conjugativo y suicida pOJ260 en los sitios de restricción *Eco*RI-*Pst*I, así se obtuvo el plásmido pME (Figura 60), que contiene la secuencia de ADN codificante de la proteína híbrida FlvA* truncada en su extremo carboxilo terminal.

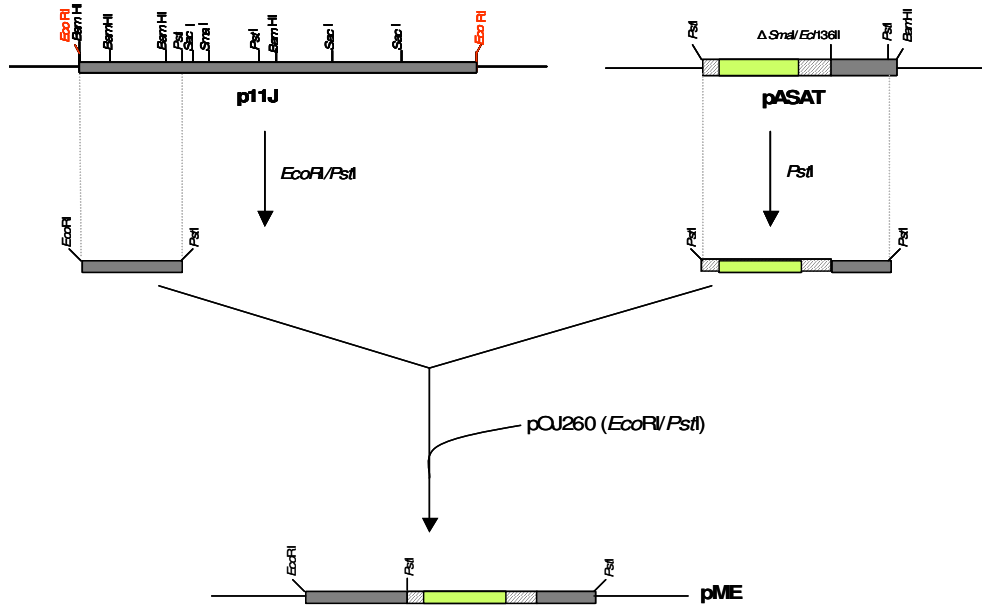


Fig. 60. Esquema de la construcción del plásmido pME, utilizado para llevar a cabo la conjugación bacteriana entre *E. coli* y *S. sp.* CO-AC25 y producir la sustitución del dominio AT de FlvA por el dominio AT de RimA.

La construcción resultante fue secuenciada con el fin de comprobar que no existía ningún cambio en el marco de lectura de la proteína híbrida generada, conservándose así una ORF intacta. En la figura 61 se muestra un alineamiento de la secuencia de aminoácidos en las regiones de unión entre RimA, FlvA y FlvA*.

	N-terminal											
RimA	RGAAEAFP	VFGEALREVC	DTLDPLLARP	LTSVMWADAD	SEEATLLHNA	EF SQPSLFAL	QVALYRLYES	WGMAPDRLAGHS				
FlvA*	MADELMDR	FPVFAELQVC	DTLDPLLARP	LTSVMWADAD	SEEATLLHNA	EF SQPSLFAL	QVALYRLYES	WGMAPDRLAGHS				
FlvA	PGMADELM	DRFPVFAELQ	ACADALTPHV	DWPVMDVLRG	RTDTPSLDRP	DVGQPALWAM	MVSLAALWRS	WGVVPAVIGSS				
		↓										
		sitio <i>Pst</i> I										
	C-terminal											
RimA	ARDPEHWVRH	VRQPVRFKDA	IGRLREERTV	GFLELGAEPL	LTPMIDECLE	AAGPQHGTAV	VPSLSSGVPD	RQILLSAAAR				
FlvA*	ARDPEHWVRH	VRQPVRFKDA	IGRLREERTV	GFLELGAEPL	LTPMIDECLE	AAGPQHGTAV	FGSADHLTAA	LAVLHAHGVL				
FlvA	HWCNLRHTA	DFESATRAAL	AAGHRLLVES	SPHPVLTRAV	QETAEDAARD	AVVLGTLRRG	FGSADHLTAA	LAVLHAHGVL				
							↓					
							sitio <i>Sma</i> I					

Fig. 61. Alineamiento de la secuencia de aminoácidos de las proteínas RimA, FlvA y la proteína híbrida FlvA*. En rojo se marcan los aminoácidos codificados por los sitios *Pst*I y *Sma*I así como el motivo conservado de cada dominio AT.

En la siguiente tabla se muestra un resumen de todas las construcciones realizadas:

Plásmido	Descripción / Uso
p11J	Plásmido de <i>E. coli</i> derivado del vector pIJ2925, que contiene un fragmento <i>EcoRI</i> de 14 kb proveniente del fago EMBL4/11J.
pAT	Plásmido de <i>E. coli</i> derivado del vector pUC18, que contiene el fragmento amplificado a partir del plásmido pSM500 con los oligonucleótidos AT-I y AT-II. Contiene el dominio AT de la proteína RimA flanqueado por los sitios de restricción <i>SacI</i> y <i>PstI</i> .
pASAT	Plásmido de <i>E. coli</i> . Contiene el fragmento de 2,5 kb <i>BamHI/SmaI</i> del plásmido p11J (dominios ACP del módulo de arranque de la proteína FlvA y los dominios KS y AT del primer módulo de elongación de la proteína FlvA) junto con el fragmento <i>Ecl136II /PstI</i> del plásmido pAT (dominio AT de la proteína RimA) clonados en el vector pUC18, digerido con <i>BamHI/PstI</i> .
pME	Plásmido de <i>E. coli</i> derivado del vector conjugativo y suicida pOJ260, que contiene el fragmento <i>PstI</i> de 2,9 kb del plásmido pASAT junto al fragmento de 4 kb <i>EcoRI/PstI</i> del plásmido p11J (dicho fragmento alberga el gen <i>orf1</i> y el extremo 5' terminal del gen <i>flvA</i>) clonados en los sitios de restricción <i>EcoRI/PstI</i> .

Tabla 14. Construcciones realizadas para llegar a obtener el plásmido pME.

Generación de FlvA* completa mediante recombinación “*in vivo*”

Obtenida la proteína FlvA* híbrida, el paso siguiente debía ser el intercambio del fragmento nativo de FlvA por el fragmento recombinante, para lo cual se decidió utilizar como estrategia la recombinación “*in vivo*” (Figura 62).

Mediante conjugación y por recombinación homóloga simple, el plásmido pME se integra dentro del cromosoma con dos alternativas posibles: bien por el extremo 5' terminal del plásmido pME, en cuyo caso la proteína que se genera es idéntica a la proteína FlvA de la cepa silvestre (Figura 62(A)); o bien por el extremo 3' terminal del plásmido pME, en cuyo caso la proteína que se generaría sería la proteína híbrida esperada, que contiene el dominio heterólogo AT (Figura 62(B)).

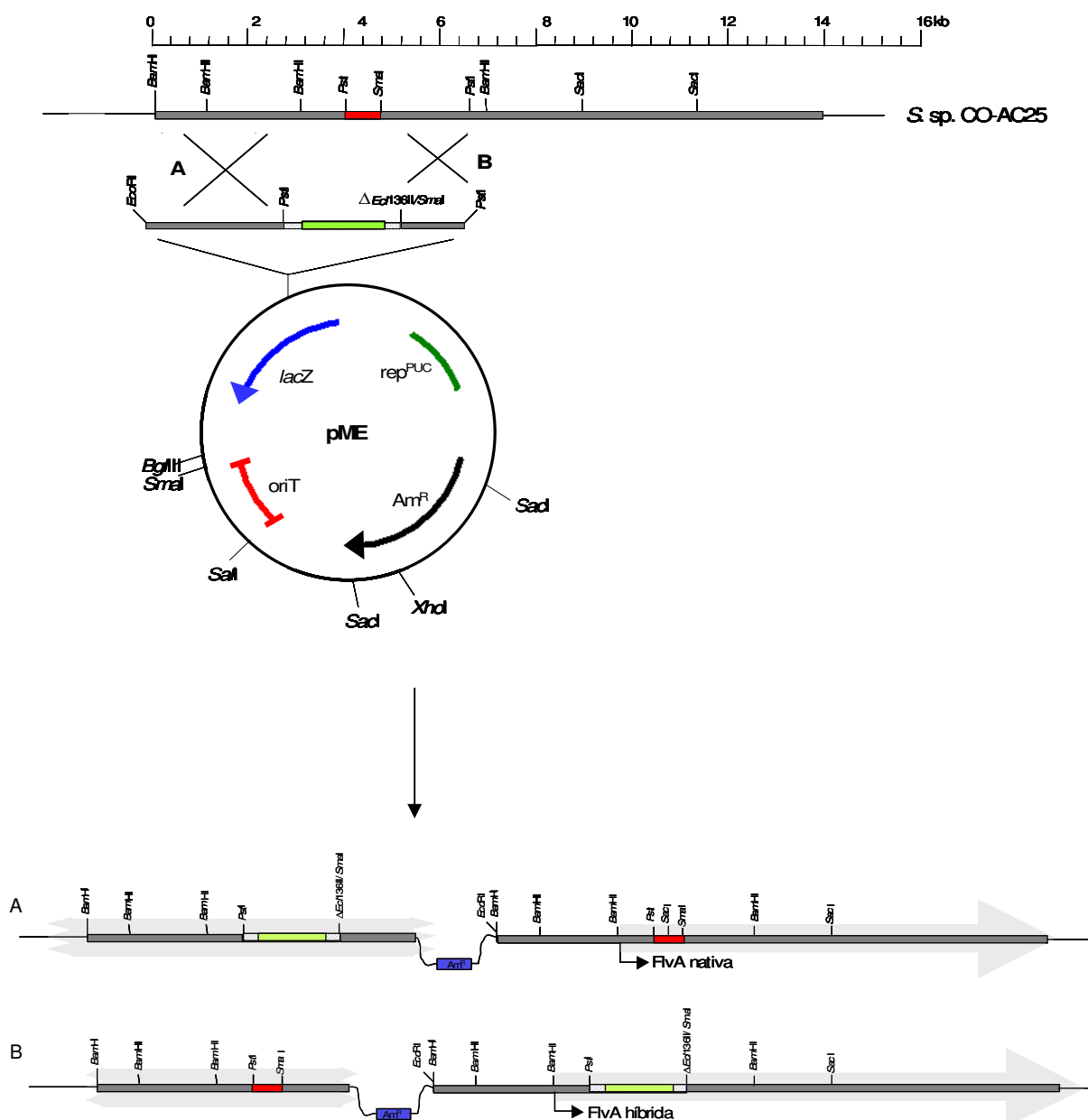


Fig. 62. Esquema de la recombinación homóloga que tiene lugar entre el plásmido pME y el cromosoma de *Streptomyces sp. CO-AC25*.

Se aisló un exconjungante resistente a apramicina al que se dio el nombre de *Streptomyces sp. CO-AC25::pME* y se realizó un *Southern Blot* con el mismo (Materiales y Métodos, apartado 2.11) con el fin de ver dónde había tenido lugar la recombinación homóloga. Como sonda se utilizó el plásmido pME marcado radiactivamente con $[\alpha^{32}\text{P}]\text{dCTP}$ (Materiales y Métodos, apartado 2.11 y 2.12) (Figura 64). Los resultados obtenidos evidenciaron que la recombinación homóloga había tenido lugar por el extremo 3' terminal del plásmido pME homólogo a la PKS de flavofungina I y que lo que se originó fue el gen que codificaba la proteína recombinante deseada.

Evaluación de la actividad biológica

Una vez confirmada la obtención del gen híbrido *flvA** se procedió a su análisis funcional.

Para ello se extrajo con MetOH una placa de R5 sólido donde se había crecido durante 5-6 días la cepa *S. sp.* CO-AC25::pME. El extracto se analizó por HPLC y los resultados pusieron de manifiesto que el exconjugante no producía ningún tipo de compuesto detectable en el rango del UV-visible, ni siquiera el macrólido oxopentaeno flavofungina I.

Se barajó la posibilidad de que este resultado se debiese a un efecto polar ejercido por la presencia del plásmido pOJ260 (presente en el plásmido pME) integrado en el cromosoma de *S. sp.* CO-AC25. Para evitar dicho efecto se procedió a seleccionar la doble recombinación con el exconjugante obtenido, de manera que se eliminara el plásmido insertado en el cromosoma. La doble recombinación tuvo lugar tras varios pases del recombinante sencillo en medio sólido sin selección con el fin de facilitar la pérdida del plásmido en el doble recombinante. Se seleccionaron varios exconjugantes que habían perdido la resistencia a apramicina.

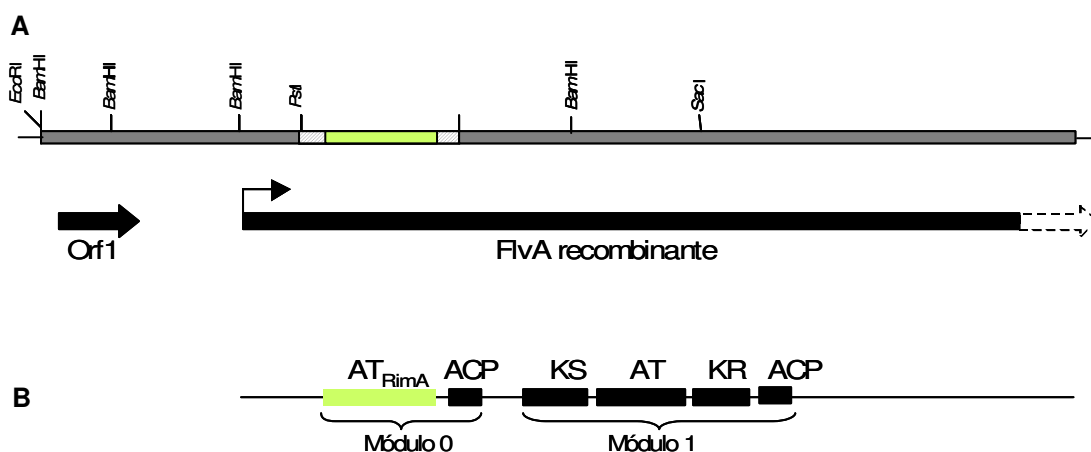


Fig. 63. A. Exconjugante *S. sp.* CO-AC25::pME originado de la doble recombinación.

B. Proteína recombinante generada. En rojo se señala el dominio AT heterólogo.

Los ADNs cromosómicos de los exconjugantes obtenidos se analizaron por *Southern Blot* utilizando el plásmido pME como sonda; así se comprobó que en todos los casos se obtenían las estructuras cromosómicas previstas. En el caso del exconjugante obtenido mediante recombinación sencilla, aparecen unas bandas adicionales, ausentes en el doble recombinante y en la cepa silvestre, que se corresponden con la presencia del plásmido pOJ260 integrado en el cromosoma (Figura 64).

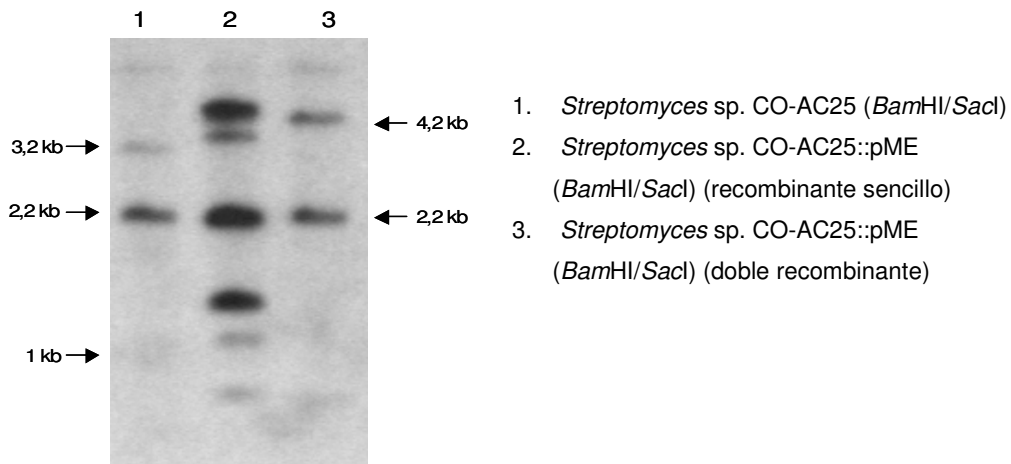


Fig. 64. *Southern Blot* que confirma la estructura cromosómica prevista en la sustitución del dominio AT de FlvA.

De igual modo que para el recombinante sencillo, se procedió a evaluar la producción del doble recombinante mediante análisis por HPLC y, al igual que ocurría en ese caso, no se detectó producción alguna (Figura 65).

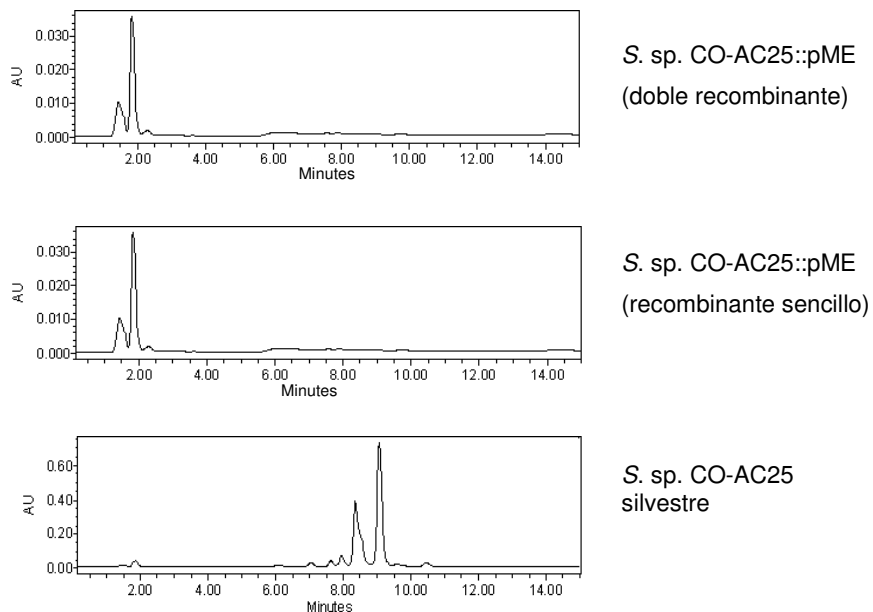


Fig. 65. Cromatogramas de producción de los exconjugantes *S. sp.* CO-AC25::pME y *S. sp.* CO-AC25 silvestre.

3.8. FUNCIONES DEDUCIDAS DE LAS REGIONES ABIERTAS DE LECTURA DE LA PKS DEL COMPUESTO BACTERICIDA

Según los datos de disrupción, la región del cromosoma aislada en los fagos recombinantes EMBL4/P1 y EMBL4/ABC a partir de la genoteca de *Streptomyces* sp. CO-AC25 se encuentra implicada en la biosíntesis del compuesto bactericida. Con el fin de poder conocer la naturaleza de este compuesto, se procedió a la caracterización parcial de los genes implicados en su formación.

Una región, de aproximadamente 24 kb, fue subclonada y parcialmente caracterizada, encontrando en ella, por el momento, tres regiones abiertas de lectura. De los tres genes que se deducen, dos se transcriben en la misma dirección mientras que el tercero se transcribe en dirección contraria (Figura 66).

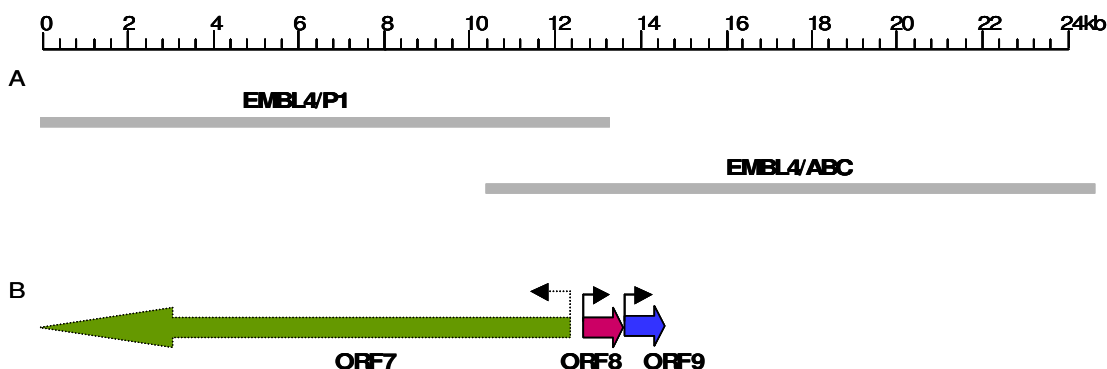


Fig 66. A. Disposición de los fagos recombinantes EMBL4/P1 y EMBL4/ABC. **B.** Genes deducidos de la región clonada.

Las tres regiones abiertas de lectura, ORF7, ORF8 y ORF9, fueron analizadas por separado:

ORF7: se trata de una región abierta de lectura incompleta de aproximadamente 13 kb parcialmente subclonada y caracterizada a partir de la región codificante para el extremo N-terminal. Comienza con un codón de inicio GTG (valina) precedido de un probable sitio de unión al ribosoma (RBS) AGGAA 15 nucleótidos corriente arriba de dicho codón.

ORF8: región abierta de lectura que comienza en un codón de inicio GTG, 510 pb corriente abajo del inicio de la ORF7 y que está precedido por el probable RBS GAGAA 4 nucleótidos corriente arriba de su inicio. Se trata de un gen de 947 pb cuyo contenido en G+C es del 69,6% y que termina en un codón de parada TGA (Anexo III).

ORF9: comienza con un codón de inicio ATG (metionina) que solapa con el codón de terminación TGA de la ORF8. Es un gen de 797 pb con un contenido en G+C del 65,6% que termina en el codón de terminación TGA (Anexo III).

3.9. CARACTERIZACIÓN DE LAS REGIONES ABIERTAS DE LECTURA DE LA PKS DEL COMPUESTO BACTERICIDA

La comparación de la secuencia de aminoácidos deducida de cada una de las regiones abiertas de lectura con proteínas de la base de datos, permitió asignar una posible función para cada una de ellas, que se muestra en la siguiente tabla:

ORF	Polipéptido	Codón de inicio / Codón de paro	RBS	Función propuesta
ORF7	Orf7	GTG/XXX	AGGAA	PKS Tipo I
ORF8	Orf8	GTG/TGA	GAGAA	ABC-transportador (union a ATP)
ORF9	Orf9	ATG/TGA		ABC-transportador (transmembrana)

Tabla 15. Funciones deducidas a partir de la ORF7, ORF8 y ORF9.

ORF7: esta región abierta de lectura codifica una proteína que presenta un alto grado de similitud con PKSs Tipo I. En ella se deduce la presencia de cuatro posibles módulos constituyendo la PKS, cada uno de los cuales presenta diferentes dominios catalíticos.

Del análisis de la proteína Orf7 se deduce un posible inicio de la PKS con un módulo en el que únicamente aparece un dominio catalítico ACP asociado a tres módulos de elongación que no están secuenciados en su totalidad y en los que se han podido distinguir diferentes dominios catalíticos colocados en el orden preciso en el que van a ser utilizados.

Así, en el módulo de elongación 1 se han identificado los dominios KS, AT y KR; en el módulo 2, se han identificado los dominios catalíticos KS, KR y ACP y en el último módulo de elongación caracterizado se han localizado los dominios KS y AT (Figura 67).

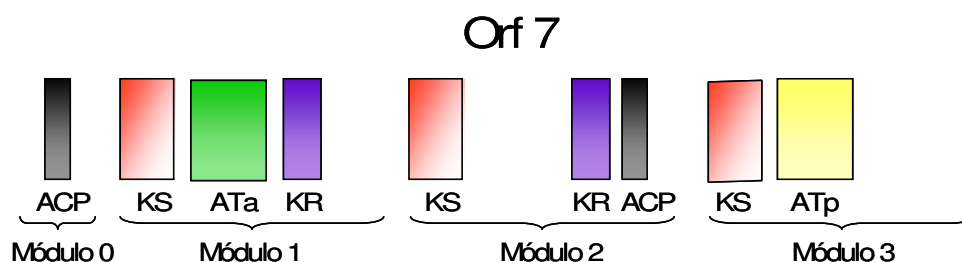


Fig. 67. Estructura modular de la proteína deducida Orf7, en la que se muestra el posible módulo de "carga" (módulo 0) asociado a tres módulos de elongación (módulos 1, 2 y 3).

El estudio de la secuencia de aminoácidos de cada uno de los dominios catalíticos muestra el grado de conservación de cada uno de ellos entre sí y con el resto de PKSs Tipo I caracterizadas:

Dominios KS: se han podido localizar los dominios KS de los tres módulos de elongación presentes en la PKS del compuesto bactericida, pero solamente en el dominio KS del tercer módulo de elongación ha sido posible encontrar la secuencia conservada característica (GPAVTVDTACSASL). Dicho dominio presenta un grado de identidad del 68% con el dominio KS de la proteína FscB de *Streptomyces* sp. FR-008 (EMBL AAQ82565) (Figura 68).

```

GPXXXXXTACSS
Orf7.1 VSVMSGRLSYAFGLEGPAVTVDTACSASLVALHLAAQALRQGECDALAGGVTVM
FscB   NSIVSGRVSYTLGLEGPAVTVDTACSSSLVALHMALQALRRGECTMALAGGVTVM
RifA   SSVASGRVSYVVLGLEGPAVTVDTACSSSLVAMHLAAQALRQGECSMALAGGVTVM

```

Fig. 68. Alineamiento de la secuencia consenso del dominio KS del primer módulo de elongación de la proteína deducida Orf7 con los dominios KS de las proteínas FscB de FR-008 y RifA de rifamicina.

Dominios AT: se han localizado los dominios AT de los módulos de elongación 1 y 3, así como los motivos conservados alrededor del residuo serina típicos de estos dominios catalíticos (GHSXG). También, para ambos módulos se han localizado las secuencias consenso que determinan la incorporación de una unidad elongadora u otra por parte de la AT. Así, la AT del primer módulo de elongación presenta la secuencia consenso RVDVV-----M-S-A-W-11aa-GHSXG responsable de que la AT incorpore propionato o metilmalonil-CoA (ATp) (Figura 69), mientras que la AT del módulo 3 presenta la secuencia consenso QTGYA-----A-FG-L-11aa-GHSXG responsable de que la unidad elongadora incorporada sea acetato o malonil-CoA (ATa) (Figura 70).

```

RVDVV-----M-S-A-W-----GHSXG
Orf7.1 NFD RVDVVQPVLFAVMVSLA VWRAGVEPAAVV GHSQGEI AAACVAGALSL
EryA1  SLE RVDVVQPVLFAVMVSLA RLWRACGAVPSAVI GHSQGEI AA AVVAGALSL
FscD   --D RVDVVQPATWAVMVSLA ELWRAHGVT PDAVI GHSQGEI AA AVVSGALT L

```

Fig. 69. Alineamiento de la secuencia consenso del dominio AT responsable de la incorporación de propionato de la proteína deducida Orf7 (módulo 1) con las proteínas EryA de eritromicina y FscD de FR-008.

```

ETGYA-----Q-A-FGLL-----GHSXG
Orf7.2 QDVMFGADAELLEQETGYAQPALFAVEVALFRLAESFGVRPEVVGGHSIGEL
AmphC REVMWGEDAGLLNETGWTQPALFAVEVALYRLVESWGVRPDFVGHSIGEI
PimS3 RDVMFG-DAEGLDETGFTQPALFAIEVALFRLAESLGVRPDFVGHSIGEI
NysC REVMWGDDVELLDETGWTQPALFAVEVALFRLVESWGVRPDFVGHSIGEI

```

Fig. 70. Alineamiento de la secuencia consenso del dominio AT responsable de la incorporación de acetato de la proteína deducida Orf7 (módulo 3) con las proteínas AmphC de anfotericina (*S. nodosus*), PimS3 de pimarcina (*S. natalensis*) y NysC de nistatina (*S. noursei*).

Dominios KR: se han localizado los dominios KR de los módulos de elongación 1 y 2, y en los dos casos se ha encontrado el motivo conservado de unión al NADP(H), siendo GXGXXGXXXA para el módulo 1 y GXGXXAXXA para el módulo 3. Ambos dominios presentan un grado de homología del 50% y 54% respectivamente con el dominio KR de la proteína PimS2 de *Streptomyces natalensis* (EMBL CAC20921).

Dominios ACP: únicamente se han localizado el dominio ACP del módulo 0 y del módulo 2 del *cluster* de biosíntesis del compuesto bactericida. En ambos casos, el porcentaje de identidad con el dominio ACP de la proteína NysC de *Streptomyces noursei* (EMBL AAF71776) es del 58% y en ambos casos, se mantiene el motivo conservado LGXD**SLXXVE** (Figura 71).

```

LGXDSLXXVE
Orf7.0 GADGTRPPGDLSRPFLDLGLDSLTAVELHRR
Orf7.3 LGYPATKTVDATRPFKDLGFDSLTAVELRNL
PimS1 LGHANPETIESTRVFQDLGFDSLTAVELRNR
AmphC LGHAAGSDIDADRAFRDLGFDSLTAVELRNR
NysC LGHADATEVETGRQFQDLGFDSLTAVELRNA

```

Fig. 71. Alineamiento de las secuencias consenso de los dominios catalíticos ACP del módulo de arranque y el tercer módulo de elongación de la proteína deducida Orf7 con las proteínas PimS1 de pimarcina (*S. natalensis*), AmphC de anfotericina (*S. nodosus*) y NysC de nistatina (*S. noursei*).

ORF8 y ORF9: codifican dos proteínas de 315 aa y 265 aa respectivamente a las que se les ha denominado Orf8 y Orf9. El peso molecular de estas proteínas es de 34 kDa para la Orf8 y de 28,2 kDa para la Orf9.

Del análisis de su secuencia de aminoácidos y por comparación con las secuencias de otras proteínas presentes en las bases de datos se ha evidenciado su homología con proteínas transportadoras que unen ATP (*ATP Binding Cassette*) o, lo que es lo mismo, con ABC-transportadores (Olano *et al.*, 1995; Fernández-Moreno *et al.*, 1998). Proteínas que son bombas de expulsión de metabolitos al exterior celular mediante el consumo de ATP.

La proteína deducida de la ORF8 presenta una identidad del 60% con la proteína Orf20 (ABC-transportador) de *Streptomyces aizunensis* (EMBL AAX98195). En la secuencia de aminoácidos de la proteína Orf8, al igual que en aquellas proteínas con las que presenta homología, se ha revelado la presencia de motivos conservados alrededor de los dominios de unión de ATP: son los motivos Walker A y Walker B (Walker *et al.*, 1982) y el lazo 3, rico en

glicina (Hyde *et al.*, 1990). En la figura 72, se muestran los alineamientos de dichos dominios con los de otras proteínas de la base de datos.

		GXXGXGK	
Orf8	<i>S. sp.</i> CO-AC25	VVDGLSFDVYRGEVFGLL	GPNGAGK STTVGMMTTRVRPTSGSVVIDG
Orf20	<i>S. aizunensis</i>	AVDDLSFSVRRGEVFGFL	GPNGAGK TTTTIGILTTRVAPTAGRAVVOG
KasK	<i>S. kasugaensis</i>	VVALDSVLSVPEGETHVLV	GPNGAGK TLVRILSTMLPTSGRALVGHV
CmrA	<i>S. griseus</i>	AVRGVDFTVAAAGEIVGFL	GPNGAGK TTTMRMLTLLRPTSGTAVIAG
		SGG	VLFLDEP
Orf8	<i>S. sp.</i> CO-AC25	DM SGG QVQRVMARALIH	RPN VLFLDEP ATGLDPQ
Orf20	<i>S. aizunensis</i>	EL SGG QAQRVMARALM	HRPD VLFLDEP ATGLDPQ
KasK	<i>S. kasugaensis</i>	GGF SRG MRQ-LHLARGL	IGPA VLFLDEP MGMDPLA
CmrA	<i>S. griseus</i>	SL SGG QRRRLDIALGL	VHQPP LIFLDEP STGLDPT

Fig. 72. Alineamiento de la proteína Orf8 de *Streptomyces sp.* CO-AC25 con las proteínas transportadoras ABC de *Streptomyces aizunensis* (Orf20), *Streptomyces griseus* (CmrA) y *Streptomyces kasugaensis* (KasK). Se muestran los motivos conservados Walker A (GXXGXGK), Walker B (VLFLDEP) y el lazo 3 (SGG).

La proteína deducida a partir de la ORF9 muestra un porcentaje de identidad del 53% con la proteína de membrana del ABC-transportador de *Streptomyces aizunensis* (EMBL AAX98194).

Orf2	VFDREFTLPITWRPAPMVGYLLGDVVRYFTASVVMITVGVVIGYRPGGATGILLGIALL
DrrB	VTDRRLRITLPVSRAAVPLGQSVADVLVTAAGTVPLLLVGLAVGWRVEGGALRAVGAALGLL
MtrB	YLERLRVTPASRLGLLLGRVLDVLAFLTIQSVVIVLIALLFGLRAP--FGGVVAATALM
Orf9	REIEDRLLSPISLSLVAVEKILFGALYGFVASLMMTPVGMILLDLSWP-ISKTIIPALGIA

Fig. 73. Alineamiento de la proteína Orf9 de *Streptomyces sp.* CO-AC25 con las proteínas Orf2 de *Streptomyces rochei*, DrrB de *Streptomyces peuceticus* y MtrB de *Streptomyces argillaceus*.

Todas estas proteínas, DrrB, MtrB,... constituyen los dominios transmembrana de los ABC-transportadores, los cuales presentan regiones hidrofóbicas mediante las cuales están embebidas en la membrana permitiendo así el transporte de metabolitos al medio extracelular.

3.10. EXPRESIÓN DE LOS GENES *orf8* Y *orf9* EN *Streptomyces lividans* TK21

Debido a la disposición de los genes que codifican la bomba de expulsión dependiente de ATP (ABC-transportador), *orf8* y *orf9*, y la región abierta de lectura ORF7, codificante de la PKS Tipo I responsable de la biosíntesis del compuesto bactericida producido por *Streptomyces sp.* CO-AC25, se pensó que dicha bomba pudiera estar implicada en la expulsión del compuesto bioactivo al medio extracelular.

Con el objeto de poder confirmar este supuesto, se procedió al clonaje de los genes *orf8* y *orf9* en el plásmido pHJL401. Para dicho clonaje, se partió del plásmido pAB6, fragmento *Bgl*II/*Eco*RI de 3,6 kb proveniente del fago recombinante EMBL4/ABC clonado en el vector pUC18, en los sitios de restricción *Bam*HI/*Eco*RI, que contiene los genes *orf8* y *orf9*. Dicho plásmido se digirió con las enzimas *Pst*I/*Eco*RI y el fragmento obtenido, que igualmente contenía los genes *orf8* y *orf9*, se clonó en el vector pHJL401 en los sitios de restricción *Pst*I/*Eco*RI. A la construcción obtenida se la denominó pABC-401 y con ella se transformaron protoplastos de la cepa *Streptomyces lividans* TK21, sensible al compuesto bactericida producido por *S. sp.* CO-AC25 (Resultados, apartado 3.1), para ver si era capaz de adquirir resistencia al mismo (Figura 74).

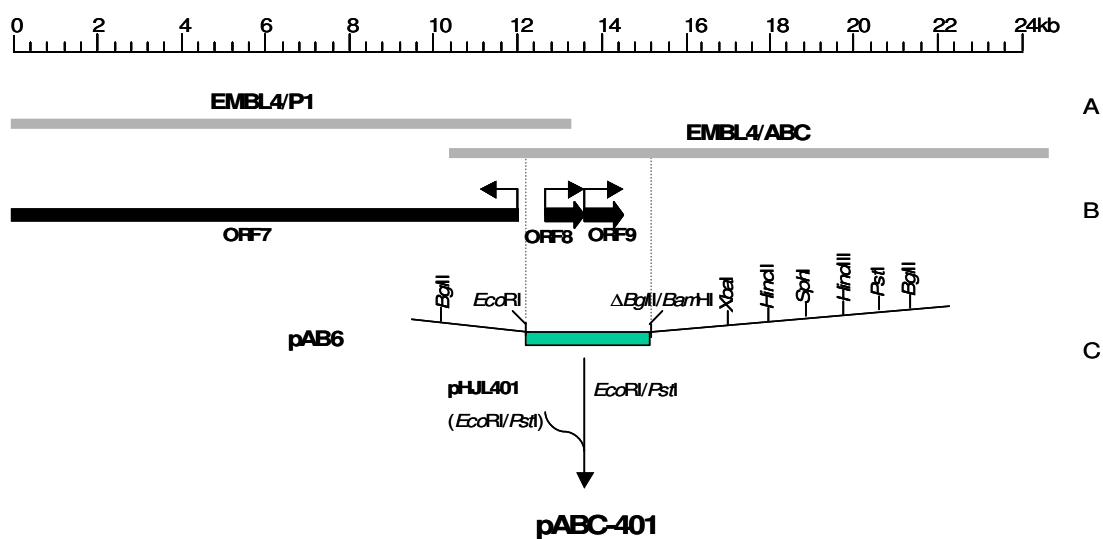


Fig. 74. Representación esquemática de la construcción pABC-401 a partir del plásmido pAB6 (C). También se muestran los fagos solapantes EMBL4/P1 y EMBL4/ABC (A) y los genes contenidos en ambos fagos: *orf7*, *orf8* y *orf9* (B).

Se obtuvieron cuatro transformantes resistentes a thioestrepton y que contenían la construcción pABC-401 (comprobado mediante la extracción de plásmido de cada uno de ellos) a los que se les denominó *S. lividans* TK21::ABC y se procedió a la realización de un ensayo de actividad biológica para ver si eran capaces de crecer en presencia del compuesto bactericida. Sobre una placa de medio sólido R5 (Materiales y Métodos, apartado 2.3) se añadió un *soft* de SNA (Materiales y Métodos, apartado 2.3) que contenía esporas de los transformantes *S. lividans* TK21::ABC y se colocaron discos de antibiograma impregnados con concentraciones crecientes del compuesto bactericida que se había extraído tal y como se indica en el apartado 2.19 de Materiales y Métodos. Como control de sensibilidad se realizó un antibiograma en iguales condiciones con *S. lividans* TK21. Tras 12 horas de incubación a 30°C, se vio que los cuatro transformantes obtenidos eran sensibles al compuesto bioactivo (Figura 75).

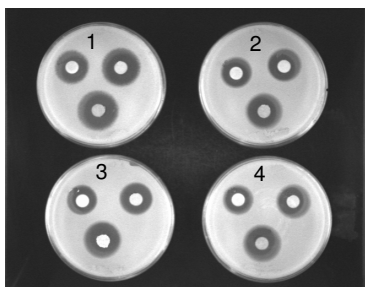


Fig. 75. Ensayos de actividad biológica del compuesto bactericida producido por *S. sp.* CO-AC 25 frente a los transformantes *S. lividans* TK21::ABC (1, 2, 3 y 4).

La ausencia de resistencia observada en los transformantes aislados se podía explicar bien porque la bomba ABC-transportador codificada por los genes *orf8* y *orf9* no estuviera implicada en la expulsión del compuesto bactericida al exterior celular o bien porque el propio promotor de dichos genes no fuera lo suficientemente fuerte como para promover la expresión de ambos en la cepa *Streptomyces lividans* TK21. Para estudiar la posible causa de esta ausencia de resistencia, se procedió a realizar los correspondientes estudios de expresión y a generar un gen recombinante bajo la expresión del promotor constitutivo *ermEp** (Schmitt-John y Engels, 1992; Bibb *et al.*, 1994).

Para ello el fragmento *Styl* de 2,1 kb proveniente del plásmido pABC-401 fue romado y clonado en el vector pUC18 en el sitio de restricción *Hind*II generándose el plásmido pAB12, que a su vez fue digerido con las enzimas de restricción *Xba*I/ *Hind*III para rescatar así los genes *orf8* y *orf9*. Por otro lado, el fragmento correspondiente al promotor *ermEp** se rescató a partir del plásmido pER-1 (Seco *et al.*, 2004), digerido con las enzimas de restricción *Xba*I/*Eco*RI; ambos fragmentos se clonaron en el vector pSU20 digerido con las enzimas de restricción *Eco*RI/*Hind*III. Así se obtuvo la construcción pSU20-ABCery, la cual se digirió con *Eco*RI/*Hind*III para liberar los genes correspondientes al ABC-transporter (*orf8* y *orf9*) junto con el promotor *ermEp**, que se clonaron en el plásmido pHJL401 digerido con *Eco*RI/*Hind*III. La construcción final con la que se llevó a cabo la transformación de los protoplastos de *Streptomyces lividans* TK21 recibió el nombre de pABCery-401 (Figura 76).

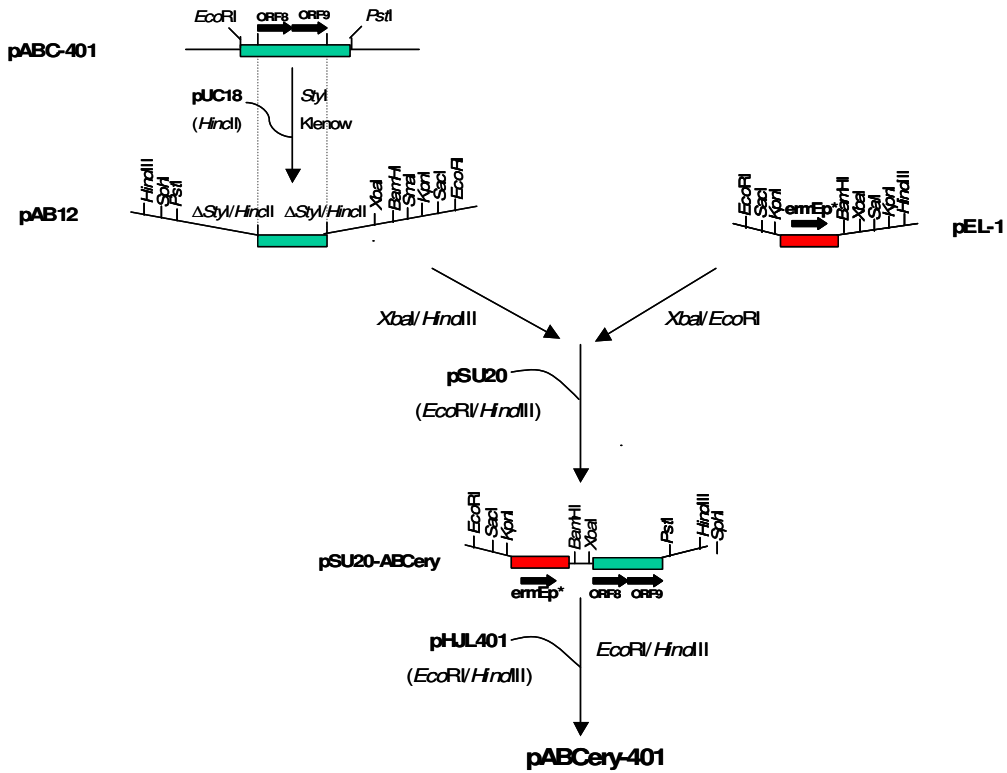


Fig. 76. Representación esquemática de la construcción del plásmido pABCery-401 a partir de los plásmidos pAB12 y pEL-1.

Al transformante obtenido se le llamó *S. lividans* TK21::ABCery y de igual forma que en el caso anterior, tras comprobar que contenía el plásmido pABCery-401, se procedió a realizar un ensayo de actividad biológica con concentraciones crecientes del compuesto bactericida. En este caso, como control se empleó el transformante *S. lividans* TK21::ABC. El resultado de dicho ensayo puso de manifiesto que el nuevo transformante obtenido seguía siendo sensible al compuesto bactericida producido por *Streptomyces* sp. CO-AC25 (Figura 77) y por tanto la bomba ABC no induce resistencia al bactericida bajo las condiciones ensayadas.

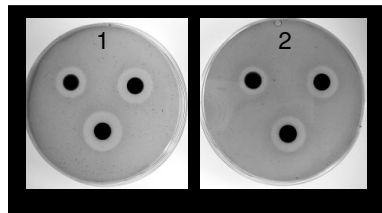


Fig. 77. Ensayos de actividad biológica del compuesto bactericida producido por *Streptomyces* sp. CO-AC25 frente a los transformantes *S. lividans* TK21::ABC (1) y *S. lividans* TK21::ABCery (2).

Estudio de la expresión de los genes *orf8* y *orf9* en *S. lividans* TK21::ABC y en *S. lividans* TK21::ABCery

Se procedió a analizar la transcripción de la bomba ABC recombinante y nativa en los transformantes *S. lividans* TK21::ABC y *S. lividans* TK21::ABCery. Para ello se extrajo el ARN de los mismos a las 48 horas de su cultivo en medio sólido R5.

Con dichos ARNs y utilizando como control el ARN de *Streptomyces* sp. CO-AC25 extraído a las 48 horas se llevó a cabo un ensayo de protección frente a endonucleasa S1 de baja resolución. Como sonda para realizar el ensayo de protección frente a endonucleasa S1 e hibridarla con los ARNs anteriormente mencionados, se utilizó el plásmido pAB5, fragmento *SacI/EcoRI* de 900 pb proveniente del cromosoma de *S. sp.* CO-AC25 clonado en el vector pUC18 en los sitios de restricción *SacI/EcoRI*. Este fragmento codifica la secuencia del extremo 5' terminal del gen *orf8* y lleva asociado una secuencia de ADN heteróloga que va a servir de control de la digestión de la endonucleasa S1 (Figura 78.A).

Los híbridos resultantes ya transferidos a una membrana de nitrocelulosa se hibridaron con a la sonda marcada radiactivamente con $\alpha^{32}\text{P}_i\text{-dGTP}$ (Materiales y Métodos, apartado 2.12) y obtenida a partir de la reacción de amplificación con los oligonucleótidos diseñados ABC-I y ABC-II y el plásmido pAB7 (fragmento *SphI/EcoRI* de 1,7 kb del cromosoma de *S. sp.* CO-AC25 clonado en el vector pUC18 en los sitios de restricción *SphI/EcoRI*) como ADN molde de la reacción de amplificación. El fragmento de PCR obtenido presentaba un tamaño de 341 pb y al igual que el fragmento proveniente del plásmido pAB5 alberga el extremo 5' del gen *orf8* (Figura 78.B).

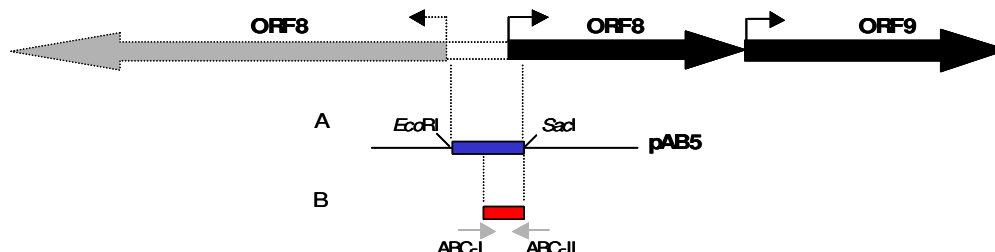


Fig. 78. A. Esquema de la sonda pAB5 utilizada en los ensayos de protección frente a endonucleasa S1. **B.** Fragmento de PCR obtenido de la amplificación del plásmido pAB7 con los oligonucleótidos diseñados ABC-I y ABC-II.

El resultado del ensayo de protección frente a endonucleasa S1 puso de manifiesto la ausencia de expresión del sistema ABC-*transporter* en el caso del transformante *S. lividans* TK21::ABC en comparación con la cepa silvestre, mientras que para el transformante *S. lividans* TK21::ABCery la expresión aumentaba bastante si se compara con la de *S. sp.* CO-AC25 (Figura 79).

Resultados

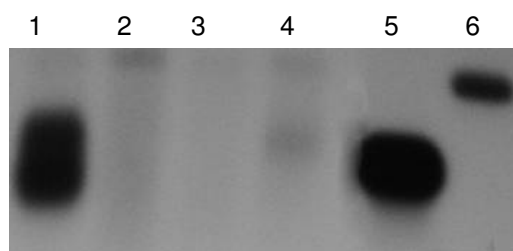


Fig. 79. Análisis transcripcional de los genes *orf8* y *orf9*. **1.** Fragmento protegido resultante para el transformante *S. lividans* TK21::ABCery. **2-3.** Análisis transcripcional para dos de los transformantes *Streptomyces lividans* TK21::ABC. **4.** Fragmento protegido para *S. sep.* CO-AC25. **5.** Fragmento amplificado utilizado como sonda radiactiva. **6.** Fragmento *SacI/EcoRI* del plásmido pAB5.

A la vista de estos resultados, queda por tanto como probable que este sistema de ABC-transportador no es capaz de inducir resistencia frente al compuesto bactericida en las condiciones ensayadas.

DISCUSIÓN

La caracterización de los metabolitos activos de una cepa de actinomiceto aislada de suelos, *Streptomyces* sp. CO-AC25, reveló la presencia de al menos dos actividades biológicas en su caldo de fermentación, por un lado una actividad antifúngica y por otro una actividad antibacteriana. Se trata de dos actividades debidas a compuestos de naturaleza química diferente que son biosintetizados por *clusters* separados. Pese a que las rutas biosintéticas de ambos están alejadas físicamente en el genoma, los datos aportados en este trabajo permiten sugerir que puedan estar interrelacionadas a nivel de regulación, ya que se ha observado que el bloqueo de la producción de uno de ellos va acompañado de un incremento en la producción del otro y viceversa. La posible existencia de una “competencia” metabólica permitiría especular que ambas rutas de biosíntesis pudieran estar compartiendo los mismos metabolitos esenciales, necesarios para el ensamblaje de las respectivas cadenas hidrocarbonadas, situación que no sólo ocurre en el organismo productor en estudio sino que ha sido descrita anteriormente para otros productores (Seco *et al.*, 2004).

La determinación de la estructura química en el caso del compuesto fungicida ha sido posible, poniendo de manifiesto que se trata de un compuesto perteneciente al grupo de los macrólidos polienos, que presenta cinco dobles enlaces conjugados con un grupo carbonilo (-C=O), lo que le caracteriza como un oxopentaeno y en cuya estructura se pueden distinguir dos regiones bien diferenciadas: una polihidroxilada y otra en la que se encuentran los dobles enlaces conjugados (R. Bógnar *et al.*, 1970). A pesar de que este tipo de polienos poseen una excelente actividad como antifúngicos, la información que hasta hoy se tiene de ellos es relativamente escasa. Varias cepas de *Streptomyces* han sido descritas como productoras de oxopentaenos, entre ellas destacan: *Streptomyces flavofungini* productora de flavofungina I y flavofungina II en proporción 10:1 (R. Bógnar *et al.*, 1970) (Figura 80), *Streptomyces ruber* (ATCC3348) productora de micoticina (Burke *et al.*, 1954; Wasserman *et al.*, 1967) en una proporción 1:1 (micoticina A y micoticina B) o *Streptomyces roseflavus* productor de roseofungina, también denominado flavomicoina (Vetlugina, 1968; Schlegel *et al.*, 1968; Hamilton-Miller, 1973). Más recientemente hay que destacar la faeriefungina producido por *Streptomyces griseus* var. *autotrophicus* (Nair *et al.*, 1989; Mulks *et al.*, 1990), mezcla inseparable de faeriefungina A y faeriefungina B en proporción 1:1, similar a flavofungina y micoticina aunque con propiedades biológicas ligeramente diferentes y RK-397, producido por *Streptomyces* sp. 87-397 (Kobinata *et al.*, 1993). A diferencia de los anteriores, *Streptomyces* sp. CO-AC25 es productora de un único oxopentaeno, flavofungina I (Figura 80) siendo indetectable la presencia de otros compuestos de este tipo.

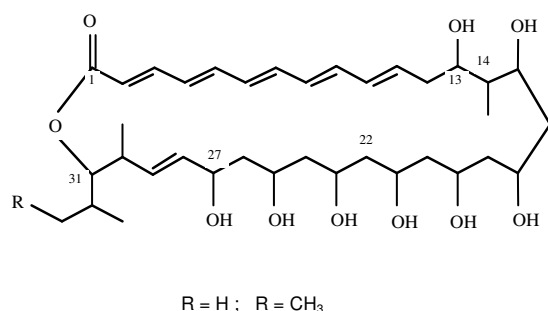


Fig. 80. Estructura química de flavofungina

I: R=H
II: R=CH₃

Comparados con los macrólidos polienos “clásicos”, como pimarcina (Aparicio *et al.*, 2000), nistatina (Brautaset *et al.*, 2000), anfotericina B (Caffrey *et al.*, 2001), candicidina (Campelo *et al.*, 2002), rimocidina y CE-108 (Seco *et al.*, 2004), los oxopentaenos presentan diferencias tanto a nivel estructural como a nivel de propiedades químico-biológicas. Se trata de compuestos muy oxidados, carentes de azúcar y de grupo carboxilo lateral unidos al anillo de macrolactona, ambos presentes en las estructuras de los polienos anteriormente mencionados. Esas diferencias estructurales hacen que los espectros de absorción de los macrólidos oxopentaenos difieran de los espectros de los macrólidos polienos “típicos”, mientras que estos últimos presentan tres máximos de absorción, los oxopentaenos presentan un único máximo y además con una longitud de onda algo desplazada hacia la derecha (Figura 81).

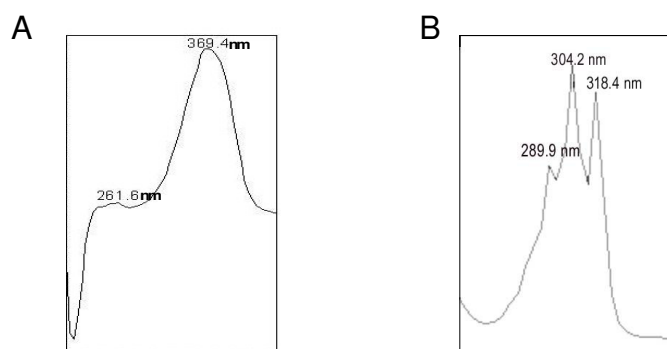


Fig. 81. A. Espectro ultravioleta característico de macrólidos oxopentaenos. B. Espectro ultravioleta característico de macrólidos polienos “típicos”.

A nivel biológico, los oxopentaenos llevan asociadas otras actividades además de las antifúngicas. Así a faeriefungina se le atribuyen propiedades antibacterianas, e incluso se han descrito actividades frente a nemátodos y larvas de mosquito (Nair *et al.*, 1989; Mulks *et al.*, 1990); a flavomicoina se le asocian propiedades antivirales; el oxopentaeno RK-397 presenta actividad frente a células tumorales (Kobinata *et al.*, 1993) y a flavofungina se le atribuyen propiedades antivirales, antitumorales e inmunosupresoras (Kino *et al.*, 1987).

La diversidad de actividades biológicas junto con las diferencias estructurales citadas anteriormente, llevan a pensar en un mecanismo de acción ligeramente distinto al de los macrólidos polienos convencionales, los cuales basan su acción en la interacción con los esteroides de las membranas celulares. Aunque tales interacciones tienen lugar también con estos compuestos, las actividades biológicas adicionales bien pudieran deberse a interacciones con otros componentes celulares o a interferencias en algún proceso metabólico. Ello, hace de estos compuestos un sistema atractivo para analizar las relaciones estructura-función en el control de procesos metabólicos, que de esa forma permitirían disponer de herramientas adicionales para posteriores manipulaciones de las rutas metabólicas de otros macrólidos.

La caracterización parcial del *cluster* biosintético de flavofungina I ha revelado la arquitectura de algunos genes implicados en su biosíntesis. Así la disposición de los dominios catalíticos de la proteína FlvA parece estar próxima a la de otras PKSs, tal como la proteína DEBS1 de eritromicina (Donadio y Katz, 1992), y en base a esta disposición, se le atribuye su implicación en el módulo de “carga” de la ruta de biosíntesis de flavofungina I. En ambas proteínas la ausencia del dominio catalítico KS en su extremo N-terminal parece apoyar esta observación y al igual que ocurre también con las rutas biosintéticas de eritromicina (Cortés *et al.*, 1990) candicidina (Campelo y Gil, 2002) o rapamicina (Schwecke *et al.*, 1995), el módulo de “carga” se encuentra asociado a varios módulos de condensación y no aislado constituyendo una única PKS a diferencia de nistatina (Brautaset *et al.*, 2000), anfotericina (Caffrey *et al.*, 2001), pimaricina (Aparicio *et al.*, 1999), rimocidina y CE-108 (Seco *et al.*, 2004). Se ha de destacar la ausencia, en este módulo, del dominio CoA-ligasa presente en las PKSs de pimaricina (Aparicio *et al.*, 1999), rimocidina y CE-108 (Seco *et al.*, 2004).

El dominio AT presente en el módulo de “carga” de la proteína FlvA da cuenta de la incorporación inicial de metilpropionil-CoA (o isobutiril-CoA) y que origina la cadena lateral inicial del anillo de macrolactona. Es de resaltar que en dicho dominio la secuencia conservada de aminoácidos entorno al residuo serina del centro activo es GSS y no GHS, a diferencia de los motivos característicos de los dominios AT de polienos “típicos”. Por su parte, los dos módulos de elongación asociados al módulo de “carga” de la proteína FlvA incorporan como unidad de elongación metilmalonil-CoA el primero (Marsden *et al.*, 1994) y malonil-CoA el segundo (Haydock *et al.*, 1995), responsables de las dos primeras condensaciones que tienen lugar durante la ruta de biosíntesis del oxopentaeno.

Los productos génicos resultantes de *flvB*, *flvC* y *flvD* son proteínas PKS Tipo I implicadas en la elongación de la cadena policetónica de flavofungina I; el orden en que actuarían está aún por determinar y necesitará de una caracterización más completa del resto de los genes biosintéticos del *cluster*. No obstante, con los datos obtenidos hasta el momento y en base a la estructura química de la flavofungina I, se ha propuesto un modelo de biosíntesis (Figura 82), en el cual se proponen las unidades incorporadas en cada etapa de elongación y el grado de reducción necesario para la formación de la cadena policetónica. El ensamblaje de dicha cadena tendría lugar a través de 15 etapas de elongación, incorporándose en todas ellas acetato (ATa) a través de la unidad elongadora malonil-CoA, excepto en el caso del primer módulo de elongación (asociado al módulo de “carga”) y en el módulo 9, dónde se incorporaría propionato (ATp) mediante decarboxilación de la unidad elongadora metilmalonil-CoA. Según este modelo de biosíntesis, las proteínas FlvB y FlvC podrían intervenir en las etapas de elongación 7 a 9 para la primera y 10 a 13 para la segunda. La liberación de la cadena policetónica del complejo PKS debería tener lugar mediante la acción del dominio funcional tioesterasa (TE) normalmente situado en el extremo C-terminal del último módulo de elongación, liberación que va acompañada de la ciclación de la cadena hidrocarbonada para finalmente originar el anillo macrocíclico de la molécula de flavofungina I. En este caso aunque no se ha identificado el dominio catalítico TE habría que resaltar una peculiaridad en el proceso de ciclación de los oxopentaenos que se refiere a la diferente disposición que presenta la

cadena de flavofungina I con respecto a la de otros macrólidos polienos y en el que podría estar implicado este dominio TE. Mientras que los macrólidos polienos “típicos” presentan los dobles enlaces conjugados próximos al inicio de las condensaciones, para la flavofungina I estos dobles enlaces aparecen en posición opuesta; hecho que podría ser el responsable de la diferente forma de ciclarse la molécula así como de la disposición espacial que adoptan los grupos hidroxilo (-OH) dentro de la misma.

La estructura de flavofungina I, a su vez, lleva implícita la ausencia de genes involucrados en modificaciones post-PKS o de “adorno” existentes en otras rutas biosintéticas, entre las que cabe destacar aquellas modificaciones llevadas a cabo por glicosiltransferasas, implicadas en transferir una molécula de azúcar a la estructura de aglicona, o las realizadas por citocromos P-450 monooxigenasas que llevan a cabo oxidaciones puntuales de las cadenas laterales metilo. Por lo tanto, no es de esperar la presencia de genes que codifiquen proteínas con estas actividades en el *cluster* de biosíntesis de flavofungina I.

Respecto a otros genes codificados en fragmentos asociados al *cluster* biosintético de flavofungina I, no es posible excluir su participación en la biosíntesis del oxopentaeno. Así la proteína Orf2, homóloga a tioesterasas Tipo II (llamadas así para diferenciarlas de las tioesterasas Tipo I) que se suelen localizar en regiones adyacentes a las PKSs de macrólidos, se le atribuye una función correctora de los posibles errores que puedan cometerse a lo largo del proceso de elongación (Kotowska *et al.*, 2002). Debido a que la proteína Orf2 precede a la PKS de flavofungina I y a que presenta los motivos conservados característicos de tioesterasas Tipo II (Resultados, apartado 3.5) es probable que esta proteína tenga función semejante en la biosíntesis del oxopentaeno. Por otro lado, la proteína Orf1 reveló ser homóloga a proteínas reguladoras de la familia TetR, y aunque no se dispone de argumentación suficiente, no se puede excluir su participación en algún proceso de regulación de la biosíntesis del oxopentaeno. En términos generales, reguladores de este tipo suelen ir asociados a regulación de procesos de resistencia a múltiples drogas, procesos de biosíntesis de antibióticos, de estrés osmótico o procesos de patogenicidad (Ramos *et al.*, 2005). Algunos ejemplos de este tipo de proteínas son ArpA, que controla la síntesis de estreptomicina en *Streptomyces griseus* (Onaka *et al.*, 1995) y TylQ, regulador de la producción del macrólido tilosina en *Streptomyces fradiae* (Stratigopoulos y Cundliffe, 2002), entre otras. Debido a la similitud de la proteína Orf1 con este tipo de represores (Figura 83) y a su proximidad con los genes implicados en la ruta biosintética de flavofungina I, esta proteína podría estar ejerciendo un papel regulador de la síntesis del macrólido oxopentaeno producido por *S. sp.* CO-AC25.

Orf1	TSARPPGRPRSGVNAAVFAATLSTVEELGYARA	TVDR	I	AAAAGIRKSTVYRRW	PSKGELI	64
AcrR	MMMARTKRKRERRERILDAALRLFAEKGYAAT	TVDE	I	AKAAGVSKGTL	YRHFPSKEDLL	60
Orf1	VDCLLDAF-----	GP	AP	LTGSSRAELMTSTIR	WVAAKIGEPGVGD	112
AcrR	LALLERALEELRALLEEALLDGAPEELEALL	L	LL	LLAYLEF	FARDDPRL	120
Orf1	AVSDPALREILSTR	LQDPYRLALQ	EALDEPENRV	LFFIDVVVGV	LLHRM-GMTGEP	171
AcrR	APELEA	ILRELLAL	LAELLERGLARGID	PGAAE	LLAALLLGLLL	180
Orf1	DVDALVDMVLP	PHFTDGRGK				190
AcrR	LAEE	LAELLLA	ALAALAS			199

Fig 83. Alineamiento de la proteína Orf1 de *Streptomyces sp.* CO-AC25 con la proteína reguladora AcrR perteneciente a la familia TetR de *E. coli* (EMBL U00734).

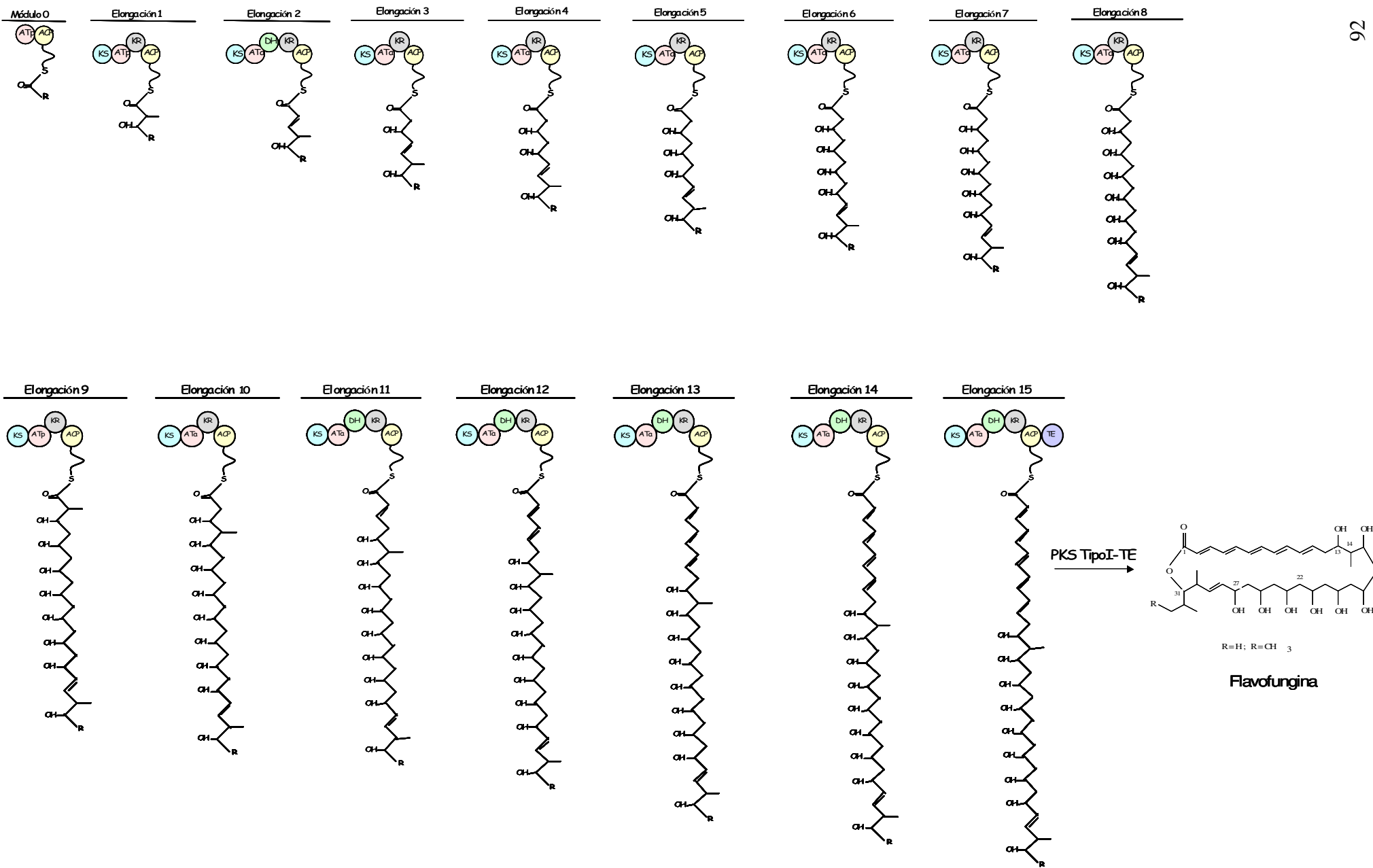


Fig. 82. Modelo propuesto para la biosíntesis del macrólido oxopentaeno flavofungina I

Tal y como ocurre en la mayoría de los *clusters* biosintéticos de antibióticos, la expresión de los genes *flv* parece estar dirigida por transcritos policistrónicos, lo que favorece que las diferentes proteínas codificadas en ellos resulten en cantidades equimoleculares. Además, llama la atención el tamaño excepcionalmente grande de los ARNs mensajeros correspondientes a estos genes que implica una alta estabilidad de los mismos. Los estudios de transcripción realizados para el *cluster* de flavofungina I revelaron la presencia de un policistrón que incluía, al menos, a los genes *flvA* y *flvB*. Debido al escaso tamaño de los espacios intergénicos entre *flvB* y *flvC* (217 pb) y entre *flvC* y *flvD* (162 pb) y a la no detección de secuencias que pudieran actuar como terminadores de la transcripción, se postula que el policistrón que codifica los genes *flvA* y *flvB* incluya también los genes *flvC* y *flvD*.

Estos ensayos, sumados a la disposición de los genes *orf1* y *orf2*, han permitido proponer un esquema de la transcripción de los genes de esta región del cromosoma de *Streptomyces* sp. CO-AC25 implicada en la biosíntesis de flavofungina I (Figura 84).

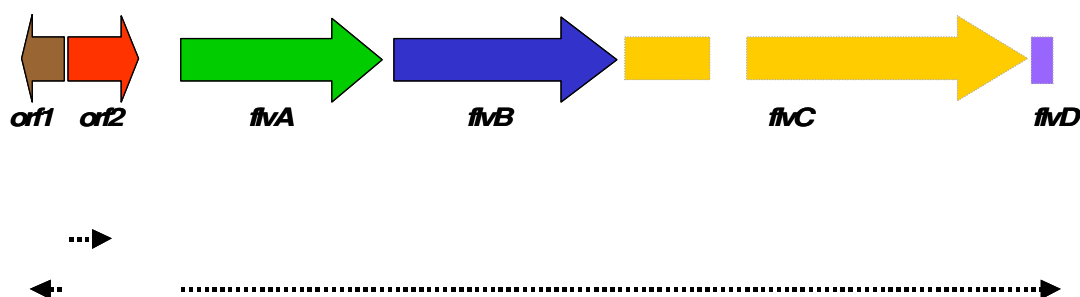


Fig. 84. Esquema de la transcripción de los genes presentes en la región caracterizada del *cluster* de flavofungina I. En flechas discontinuas se indican los ARNm esperados.

Los análisis realizados han puesto de manifiesto que los transcritos dejan de poder ser detectados aún antes de que los niveles de flavofungina I sigan aumentando, lo que permite sugerir que la producción del oxopentaeno en las condiciones ensayadas es acumulativa, y probablemente, el “recambio” de las proteínas implicadas en la producción sea de una vida media alta; este hecho, permitiría a la célula seguir produciendo el metabolito aún después de que los genes implicados en su producción hayan dejado de transcribirse.

Resulta sorprendente la no detección de metabolitos relacionados con flavofungina I una vez sustituido en el genoma un gen nativo por el híbrido correspondiente. Este hecho contrasta significativamente con los resultados obtenidos al realizar sustituciones similares en otros sistemas tales como el de eritromicina, donde la sustitución de dominios catalíticos AT presentes en distintos módulos de la PKS de eritromicina por dominios catalíticos AT de las PKSs de *Streptomyces venezuelae* o *Streptomyces hygroscopicus* dio lugar a la generación de moléculas derivadas en las que se produjeron cambios estructurales a nivel de las cadenas laterales (Ruan *et al.*, 1997; Stassi *et al.*, 1998).

El dominio AT correspondiente al módulo de “carga” de la PKS de rimocidina y CE-108, utilizado para la sustitución en *S. sp.* CO-AC25, es capaz de incorporar tanto malonil-CoA como butiril-CoA en *S. diastaticus* var. 108 (Seco *et al.*, 2004) por lo que era previsible que en

la cepa productora de flavofungina I se generase un derivado con una cadena lateral inicial más corta que la original. Aunque el intercambio de un dominio AT nativo por uno heterólogo tuvo lugar en el genoma, la cepa recombinante obtenida presentaba la ruta biosintética de flavofungina I interrumpida. Esta ausencia de producción por parte de la cepa modificada *Streptomyces* sp. CO-AC25::pME es difícil de interpretar y aunque podría deberse a que la nueva proteína recombinante generada haya perdido la funcionalidad debido, por ejemplo, a que esté sufriendo algún tipo de distorsión (física, etc) que la haga inoperante (Stassi *et al.*, 1998) se puede especular que, aún siendo funcional, los dominios AT correspondientes a los módulos de las condensaciones sucesivas sean tan estrictos en el reconocimiento de los grupos acilo que sean incapaces de originar las condensaciones requeridas. Con este planteamiento en mente, sería necesario abordar la manipulación de los dominios AT posteriores para reducir la especificidad en el reconocimiento de los grupos acilos. Resultados similares se han obtenido al intercambiar el dominio AT del módulo de “carga” de biosíntesis de nistatina por el de rimocidina (S. Zotchev, comunicación personal). También pudiera ser que la especificidad por la unidad iniciadora metilpropionil-CoA (isobutiril-CoA) por parte del módulo de “carga” de *Streptomyces* sp. CO-AC25 no venga dada únicamente por el dominio catalítico AT y que este alto grado de especificidad todavía existente impida la incorporación de unidades iniciadoras alternativas. Finalmente, pudiera ser que la cepa modificada fuera capaz de sintetizar una estructura macrocíclica derivada de flavofungina I, pero que debido a los cambios estructurales producidos en la misma, dicha estructura fuera inestable.

La caracterización de la región cromosómica perteneciente al *cluster* implicado en la biosíntesis del compuesto bactericida reveló la presencia de una proteína, Orf7, homóloga a PKSs Tipo I, donde se distinguen al menos cuatro módulos de elongación y donde es de resaltar la presencia de un dominio catalítico ACP en el extremo N-terminal. Esta disposición podría sugerir que esa región de la PKS estuviese implicada en un posible módulo de “carga” con una disposición poco convencional. No obstante cabe la posibilidad de que dicho dominio no fuese funcional; sería necesario hacer nuevas aproximaciones para disponer de datos relevantes sobre su operatividad. Aunque está claro el carácter modular de la PKS implicada en la biosíntesis del compuesto bactericida, no se dispone de datos suficientes para poder especular sobre su naturaleza química. La interferencia entre el mutante en esta región y la producción de flavofungina I permiten hipotetizar que el compuesto bactericida pudiera estar biosintetizado utilizando unidades metabólicas equivalentes a las del fungicida.

Cabe destacar el hecho de que el sistema ABC-transportador codificado por los genes *orf8* y *orf9*, sea incapaz de inducir resistencia frente al compuesto bactericida. La disposición física en *cluster* de este sistema junto a la PKS correspondiente, así como la disposición divergente de la posible región promotora parece apuntar a su implicación en la excreción del metabolito al exterior celular. Se trata, en este caso, de un ABC-transportador Tipo I (Méndez y Salas, 1998), donde el gen *orf8* codificaría el componente hidrofílico de unión a ATP y el gen

orf9 codificaría el componente hidrofóbico de unión a membrana citoplasmática. Estos sistemas se encargan de la secreción de diferentes moléculas (azúcares, aminoácidos, oligopéptidos, drogas, etc.) al exterior celular utilizando como fuente de energía la hidrólisis de ATP mediante un proceso en contra de gradiente de concentración. De esta forma se evita la acumulación de sustancias tóxicas en el interior celular y las consecuencias que dicha acumulación lleva asociada.

Los genes implicados en codificar estos sistemas de transporte muchas veces se encuentran formando parte del *cluster* de biosíntesis de los compuestos que presentan actividad biocida. Por este motivo, se planteó la posibilidad de que pudiera tratarse de la proteína ABC encargada de expulsar dicho compuesto al exterior celular. Así también ocurre, por ejemplo, con las proteínas OleC-Orf5, implicadas en exportar oleandomicina en *Streptomyces antibioticus* (Rodríguez *et al.*, 1993) o las proteínas DrrA-DrrB encargadas de exportar daunorrubicina y doxorrubicina al medio extracelular en *Streptomyces peuceticus* (Guilfoile y Hutchinson, 1991), siendo frecuente también en los *cluster* de anfotericina (Caffrey *et al.*, 2001), nistatina (Brautaset *et al.*, 2000), pimaricina (Aparicio *et al.*, 2000) y candidina (Campelo *et al.*, 2002). En este caso, dada la proximidad del sistema ABC-transportador a la PKS responsable de la biosíntesis del bactericida, es probable que no esté implicado en el movimiento del oxopentaeno u otro polieno al exterior celular. No obstante, tampoco se puede descartar; la disposición del promotor divergente parece sugerir un tipo de regulación coordinada con la expresión de la PKS en la que un posible regulador pudiera activar el promotor en una dirección en tanto reprimiese la expresión del gen de dirección opuesta. Si este fuera el caso, el fenómeno podría correlacionarse con la activación observada de la síntesis de flavofungina I en el mutante interrumpido en la PKS del bactericida. Son necesarios nuevos ensayos para conocer con más detalle el tipo de regulación que tiene lugar en la expresión de la PKS del bactericida y este sistema ABC-transportador.

CONCLUSIONES

1. Se ha caracterizado una cepa aislada de suelos, *Streptomyces* sp. CO-AC25, productora de dos compuestos biocidas cuyo mecanismo biosintético está basado en poliquétido sintetasas (PKSs) Tipo I: un fungicida perteneciente al grupo de los macrólidos oxopentaenos, flavofungina I y un bactericida de estructura química no determinada. Entre ambas rutas metabólicas existe una “competencia” por los metabolitos requeridos para la formación de los compuestos finales, lo que sugiere un tipo de regulación coordinada.
2. En base a la estructura química del compuesto oxopentaeno, se ha propuesto un posible mecanismo biosintético que comenzaría con la incorporación de metilpropionil-CoA (isobutiril-CoA) como unidad iniciadora. La caracterización del gen *flvA* y la secuencia deducida de su producto génico parecen confirmar esta hipótesis.
3. Los resultados obtenidos de la sustitución del dominio catalítico AT correspondiente al módulo de “carga” de la proteína FlvA por un dominio catalítico heterólogo perteneciente al módulo de “carga” de rimocidina y CE-108, sugieren que los dominios AT de los sucesivos módulos de elongación son altamente específicos y determinantes en el reconocimiento de otros sustratos.
4. Se ha identificado un sistema ABC-transportador Tipo I cuyo sustrato no ha podido ser determinado. Aunque aparentemente no reconoce al compuesto bactericida no es descartable su participación en la excreción de los metabolitos producidos por *Streptomyces* sp. CO-AC25 correspondientes al bactericida ni al oxopentaeno.

ANEXOS



ANEXO I

Secuencia de nucleótidos de 35323 pb aislada del cromosoma de *Streptomyces* sp. CO-AC25 que contiene los genes *orf1*, *orf2*, *flvA*, *flvB* y el extremo 5' terminal del gen *flvC*. Se muestra la secuencia de aminoácidos deducida de cada una de las regiones abiertas de lectura. En negrita se marca la secuencia de nucleótidos codificante. Los inicios de transcripción de cada uno de los genes se indica con flechas, discontinuas cuando el gen está incompleto.

Anexo I

ATCACGGATCTGAACGGGGACAAGCAGCCCGAGATCGCCCTCCCGGTCCGGGCGAACACATCCGAGGTGTGGTTCGGTCCGGCAGGACGGCAAGATCG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 100
TAGTGCCTAGACTTGCCCTGTTCTCGTGGGCTCTAGCGGGAGGGCCAGGCCGCTTGTGTAGGCTCCACACCAGCCAGCCGTGCTTCTGCGCTTCTAGC

CCGGGGCCGCGCCACCCCGTGGTGAAGGCTCCGGTGGGCTGCTCAGGTGCCGTGGGCCGGGACGAGTTCCGGGTGTGGCGCGCTGACGGTCCCGTA
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 200
GGCCCCGGCGGGTGGGCGCACCATTCCGAGGCCACCCGACGAGTCCACGGCACCCGGCCCTGTCTAAGCCCCACAACCGCGGACTGCCAGGGCAT

GCCGCTCGTACCCGTTTCATGCCCCAGGGGCCCTGTCGCTCGACTCGTCCGGCAGGGCTCCTGGGGCATCTCCATGTCTGGGGCGGGTGTCCGGTAC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 300
CGGCGAGCATGGGCAAGTACCGGGTCCCGGGACAGCGAGCTGAGCAGCCGCTCCCGAGGACCCCGTAGAGGTACAGACCCCGCCACAGCCACTG

GGGCGCCGGATGTCTCCCGACGTCGGATGTCCCGCCGGGCGGACGTTGCCCGACGGGGCATCACATCTTGCCGCGCCGTCGGTGAAGTCCGGCAGGA
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 400
CCCGCGCCTACAGAGGGTGCAGCCTACAGGGGGCGGCCCGCTGCACCGGGTGCCTCCGT**AGTGTAGAACGGGCCGGCAGCCACTTCACGCCGTCCT**
* M K G R G D T F H P L

CCATGTCCACCAGGGCGTCGACATCGGCCTCGACCATCGGCTCGCCGTCATGCCATCCGGTGCAGCAGCAGCCGACCACGACGTCGATGAAGAACAG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 500
GGTACAGGTGGTCCCGCAGCTGTAGCCGGAGCTGGTAGCCGAGCGGCCAGTACGGGTAGGCCACGTCGTCGTCGGGCTGGTGTGCAGCTACTTCTTGTG
V M D V L A D V D A E V M P E G T M G M R H L L V G V V V D I F F L

GACCCGGTTCCTCCGGCTCGTCCAGCGCCTCCTGGAGCGCCAGCCGGTACGGGTC**PstI**
CTGCAGCCCGGTGGACAGGATCTCCCGCAGGGCCGGGTCGGTACG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 600
CTGGGCCAAGAGGCCGAGCAGTTCGCGGAGGACCTCGCGGTCCGCCATGCCAGGACGTCGGCGCACCTGTCTAGAGGGCGTCCCGCCAGCGACTGC
V R N E P E D L A E Q L A L R Y P D Q L R T S L I E R L A P D S V

CGCTCGCTGAACACGCGCGGAACCGCTCGCCACCCCGGCTCACCGATCTTCGCCGCGACCCAGCGGATGGTTCGAGGTATCAGCTCGGCGGACTGG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 700
CGCAGCGACTTGTGCGGCCGCTTGGCGAGCGGGTGGGGCCGAGTGGCTAGAAGCGGCGCTGGGTTCGCTTACCAGCTCCAGTAGTCGAGCCGCGCTGACC
A D S F V G A F A D G V G P E G I K A A V W R I T S T M L E A R S

AGCCCGTCAGCGGCGCCGGGCGAACGCGTCCAGGAGGCAGTCCACGATCAGCTCGCCCTTCGACGCGCAGCGCCGGTAGACCGTTCGACTTCG**BamHI**
GGATCCC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 800
TCGGGCAGTCGCGCGGCCCGGCTTGGCGAGGTCCTCCGTCAGGTGCTAGTCGAGCGGGAAGCTGCCGGTCCGCGCCATCTGGCAGCTGAAGCCCTAGGG
S G T L P A P G F A D L L C D V I L E G K S P W R R Y V T S K R I G

GGCCGCGCGGCGATGCGATCGACGGTGGCGCGCGTAGCCGAGTTCCTCGACCGTGTCTCAACGTGGCGGGAAGACAGCGGCGTTGACCCGGAGCGC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 900
CCGGCGCCCGCTACGCTAGCTGCCACCGCGCGCATCGGCTCAAGGAGCTGGCACGAGTTGCACCGCCGCTTCTGTCCGCGCAACTGCGGCCCTCGCG
A A A A I R D V T A R A Y G L E E V T S L T A A F V A A N V G S R

GGACGGCCGGGCGCCTGGCGGATGTGGTTCGATGTATAGCCGACGATACCTCTTTTCGCTACCTTGAGTATCGATACTGAAGTCGCGAAACTCTCTC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 1000
CCTGCCGGCCCGGACCGCCTACACCAGCTACAGT**TCGGCGTGTAT**GGAGAAAAG**CGATGGAACCTATAGCTATGGACTTCAGCGCTTTGAGAGAG**
P R G P P R A S T T S T M
orf1 ←

Anexo I

GCGGCGGGCTCGCGAACGAGCCGGGCAACGAGCTGGCCAACGGGCTGACGCATGAGCTGGTCAACGAGCCACCCACGAACCTCACTGGCGACGAGACCCC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 1100
CGCCGCCGAGCGCTTGTCTCGGCCGTTGCTCGACCGGTTGCCGACTGCGTACTCGACCAGTTGCTCGGGTGGGTGCTTGAGTGACCCTGCTCTGGGG
TTTGAGC **RBS**
AAGGAGAAGGCATCAGCGAGAGCACCACCGCGGAGCGGGCCCGGAGCCCGACTGGCTGGTGGATGACCACCGAGGAGCTGGTGCCTACC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 1200
AAACTCGTTCTCTTCCGTACTCGCTCTCGTGGTGGCGCCTCGCCGGCGGCTCGGGCTGACCGACCACTACTGGTGGCTCCTCGACCAGCGGATGGCGC
M S E S T T A E R P P E P D W L V M T T E E L V A Y R D
└─▶ **orf2**
ACGCGGAGAACCGCTTCCGGGCGTCCAGCGGATCCGGGCGATGACCGGGAGCCGACCCCGCGCCACGATCCGATGGGAGAAGGTGGCGCTGCCCGG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 1300
TGCGCTCTTGGCGAAGGCCCGCAGGTGCGCTAGGCCCGCTACTGGCCCTCGGCTGGGGCCGCGGTGCTAGGCTACCCTCTTCCACCGCGACGGGCC
A E N R F R A S S A I R A M T G E P D P G A T I R W E K V A L P G
CCGTGAACCTCCGGTCCGGGTGCACCGACCGGTACCGGGCGGACGGCGAAGGCGTGCAGCGGGGGGGCTGCCGTCGTCATCCACGTGCACGGGGC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 1400
GGCACTTGAAGGCCAGGCCACGTGGCTGGCCAGTGGCCCGGCTGCCGCTTCCGCACGCCGCCCGCCCGCCGACGGCGAGCAGTAGGTGCACGTGCCCCCG
R E L P V R V H R P V T G R D G E G V R R G G L P L V I H V H G G
GGGTCGTGGGTACGGCCGTGCAGTGCAGTGGGTACCGTCCAGTCCAGTGGCCGCCGACTGCCCGCCGTCGTCGCTCTCGGTGGAGCACCGCCTCGTCGACC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 1500
CCCAAGCACCCATGCCGGCACGTACCGTGCACCCAGTGGTGCAGTGCACCGGGGGTGCAGGGCGGCAGCAGCAGAGCCACCTCGTGGCGGAGCAGCTGG
G F V G T A V Q C D W V T S H V A A R L P A V V V S V E H R L V D P
CGGACCATCCGCTGTCCGGCCCGTGCAGCAGCGGTGGGACGTGCTCCAGCAGTGGTGCAGCACCGCCGCGCAGTGGGGCGTGGACCCGGAGCGGACCGC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 1600
GCCTGGTAGGCGACAGCCGGCGGCGAGCTGCTGCCGACCCTGCACGAGGTCGTGCACCACGCCGTGCCGCGGTCACCCCGCACCTGGGCTCGCTGGCG
D H P L S A A V D D G W D V L Q H V V R H A A Q W G V D P E R T A
CGTCTTCGGCGAGAGCTGCGGCGCGTGGATCAGCGCCCTGGCGGCGTCCGGGCCAGGGAGGCCGGCCTGGGGCTGCGGGCGCAGGTGCTGGTGAACCC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 1700
GCAGAAGCCGCTCTCGACGCCGCGCAACTAGTCGCGGGACCGCCGCGAGGCCCGGTCCCTCCGGCCGACCCCGACGCCCGCTCCAGCACCCTGGGG
V F G E S C G A L I S A L A A L R A R E A G L G L R A Q V L V N P
GCTGTCGACGTGACCGGGCGATGCTCGACCACGCCCGGCGCCGAGCACGCGCACACCCCGACCTCGCCGTGCCGACGTGCGACTCTTCCAACGGC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 1800
CGACAGCTGCACTGGCCCCGCTACGAGCTGGTGGGAGCCGGCGGCTCGTGCAGGTCGTGGGGCTGGGAGCGGCACGGCGTCGACGCTGAGAAGGTTGCCG
A V D V T G A M L D H A S A A E H A H T P T L A V P Q L R L F Q R L
TCGCGGTTCCCGGGGATCCGACCCTCGCGGCTCTCGCCGCTGTACGCCGACACTCACCGGGCTTGCCCCGGCGCTCGTGGTGGTGGCGACCCACGA
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 1900
AGCGCAAGGGGGCCCTAGGCTGGGAGCGCGAGAGCGGGACATGCGGCTGCTGGAGTGGCCGAACGGGGCCGCGAGCACCACCGCTGGGTGCT
A V P P G S D P R A L S P L Y A D D L T G L A P A L V V V P T H D
CCCGTTGGCCGACCACGGCCCGCTACGCCGAACGACTGCGGGCGGCGGGACGCCCGTGGCGCTCACCGAACACGCCGAAGCGCCGACCGGTTCCCT
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 2000
GGGCAACCGGCTGGTCCCGCGGCGATGCGGCTTGTGACGCCCGCGGCCCTGCGGGCACGCCGAGTGGCTTGTGCGGCTTCGCGCGTGCAGGAG
P L A D H G R R Y A E R L R A A G T P V R L T E H A E A P H A F L

Anexo I

AGCCTGCCGGGCGTGGTCCCGCAGGCCGCGGCCGCGCGCGGAGATCTTCGACTTCCCTCCGGGCTTCCCTCGCCGACCTGGGCTAGCCCCGGGCCCCGA
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 2100
TCGGACGGCCCCGACCAGGGCGTCCGGCGCCGGCGCGCCCTCTAGAAGCTGAAGGAGGCCGAAGGGAGCGGCTGGACCCGATCGGGCTCCGGGCG
S L P G V V P Q A A A A R A E I F D F L R A S L A D L G *
ATCCGCCGCCCGTGAGGTGACCGCAACTCGCGGGCGGGCGCCGTCTTGGGCCGACGCCACCTTCTTACGCCTTCTACCAGGGCGTCACCTGCG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 2200
TAGGCGGGCGGGCACTCCACTGGCGGTTGAGCGCCCCGCGCCGGCAAGAACC CGCCTGCGGTGGAAGAAGTGGGAGAAGTGGTCCCGCAGTGGACGC
CGACAGGGTGACGCCCTGTCCGCGCTGGGCACGTTGTGAGATGAACCACGATGATCGCGCAGGAGGGCTGAAACGGGCAAGAAGCGGAGCGCTGGTGTG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 2300
GCTGTCCCACTGCGGGACAGGCGCGCACCCGTGCAACTCTACTTGGTGCTACTAGCGCTCCTCCCGACTTGGCCGTTCTTCGCCTCGCGACCACAC
GCCAAACGGGCACGACCTGGTCCGGATGTGCCGTTTTCGTGAGTGTGCCGGTGGTGTGTGTCCGGTGGGTGTGCGGATGCGAACTGGGTG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 2400
CGGTTTGCCCGTGTGGACAGGCCTACACGGGCAAAGCAGCTCACGACGGGCCACGAGCACACACAGGCCACAGCCACAGCGCCTACGCTTGACCCAC
GAGTTGATGCCAGGTAGCGGAGCGTGTGGCTCTTCGCTTCGGCATATATAGAGGGCGCTCGGGTGCCGGGCGAGCTTCTACGCGCCAGTAAGCGA
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 2500
CTCAACTACGGTCCATCGCCTCCGCAACACCAGAGAAGCGAAGCGTATATATCTCCCGCAGCGCCACGGCCCGTGTCAAGAATGCGCGGTCCATTGCT
GCGCTGCCCGCCCGGATAGCCACAAATGACGGGGTGGTGGGCAAAGCAGGTCGATACTGTTGCGCCAGGCACGGACCAGTACATCGCATTCA
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 2600
CGCGACGGCGCGGGCTATCGGTGTTTACCTGCCCCACCAGCCCGTTTCCGTCCAGGCTATGACAACGCGGTCCGTGCCTGGTCACTGTAGCGTAAGT
CCTGACCATGAGCGATCCACCCGCTATCCGATCGGGCACCCCGGAGGAGCATTCACTTTGCCCTCCCGCATCGAGCCCGAAGCCCATATGGAGTCCAC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 2700
GGACTGGTACTCGCTAGGTGGGCGCATAGGCTAGCCCGTGGGGCCCTCCTCGTAAGTGAACGGAGGGCGTAGCTCGGGCTTCGGGTATACCTCAGGTG
CCAAAGGGAGAGCACGGCCAGTGACCCGCGATGCCGTTCTGGTACTTCCGAGCGAGCCTTCCCGGTGCCGTGTGGAGACCGGGGCCGGTTGCTGTC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 2800
GGTTTCCCTCTCGTGCCGGTCACTGGGCGCTACGGCCAAGACCATGAAGGCTCGCTCGGAAGGGCCGACGGCACACCTCTGGCCCCGGCCCAACGACAG
CCTACTCAAAGTAGGTAGTACTGGCGTCACGCGCACTTGGCAGATGGAGTGGCGCCGCCATTTTTGCTTACCCCGACGTATTGATTACCTCTGCC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 2900
GGATGAGTTTTTCATCCATCACTGACCGCAGTGCGCCGTGAACCGTCTACCTCACGCGCGGGTAAAAACGAATGGGGCTGCATAACTAATGGAGACGG
CGACCGCTGTCTTGTGAGTTCGAAACAGAGTTTTAAACAATACCTTTCCGCGTGTCCCGGGTAATGAACTCGCCATGGTTGCGCTGTCTGGCGCAC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 3000
GCTGGCGACAGAACAGCTCAAGCTTTGTCTCAAATGTTATGGAAAGGCGCACAGGGCGCCATTACTTGAGCGGGTACCAACGCGACAGGACCGCGTG
CGCTTCGTCCGGACTGGCAGTGGCGCGACGCCTGTATCCGGCCCCGGCGCCCTAGGCCGAGCACCCGTGCCATGGCCCGGTGCTGAGCCGCGGGTG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 3100
CGAAGCAGCCGCTGACCGTCAGCCGCGCTGCGGACATAGCCGGGCGCCCGGATCCGGCGTCTGGGCACGGTACCGGGCCAGCACTCGGCGCCAC
GCGCCCGGGAGTCTTTGTGCTCCAGCTCACGGTGGTGGCCGGACCTCGCGCCGTGGTGTCTCCGGGACGGCCGCTGCCGTGTTCCAGGCCAGGTCAT
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 3200
CGCGGCGCCCTCAGAAACAGCAGGGTGGAGTCCACCACGGGCTGGAGCGCGGCACCACAGAGGCCCGTCCGGCGACGGCACAAGGTCGGGTCCAGTA
CGGTGTTCCCGCGGTAACGTCACGGGTGTGCTCGGCGCTTTCCGCTACGCGCGGGTGGGTGTTTCGACCGGATGCGGATAGGTGTCCCATGTTAAG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 3400
GCCACAAGGGCCGCAATTGCAAGTCCACACGAGCCGCAAAGGCGATGCGCGGCCACCCACAAAGTGCCTACGCTATCCACAGGGGTACAATTC

Anexo I

GGATGCGTGCACCCGAGGTTGACGGTGCACGGCAGTCGGGCATCGAGAGGGGTACCGGATGTTCCCTTCCGCTTATAGCGGTGTGCCCTTTCAGTTG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 3500
CCTACGCACGCTGGGCTCCAAGTCCAGCGTGCCTGAGCCGTAGCTCTCCCATGGCCTACAAGGGGAAGGCGAATATCGCCACACGGGAAAGTCAAC

CCCCAAAGGGCAGCGGAGTGGCTGCGTCCAGCGGGTGAAGCTGGGTTTTGCGTACCCTGGCAGCGCCGGGGCGGGTTAATTCCTCAAGTCGGCCCC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 3600
GGGTTTTCCGTCGCTCACCAGCGAGTGCCTTCTCGACCCAAAACGCATGGGACCGTCGCGGCCCGCCCAATTAAGGAGTTCAGCCGGGG

GTAAATTCGTGAGCCCGGGCGCCGACTCGGCCAACACGGCGCAACGGGATTCTCTCGCGGGCGCGTTACGACACGTTGTCCGACACGCGGCAG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 3700
CATTTAAAGCACTCGGGCCCGCGGGCTGAGCGCGGTTGTGCCGCGTTGCCATAAGGAGCGCGCCGCGCAATGCTGTGCAACAGGCTGTGCGCCGTC

ACTGGAAATAGCCGTCGACCGACCGAAATATACGCCAGTTGACAATTTCTGGAGCGCTGGTCTTGACTTGCGCCAGTTGTGGTGCCTTTGGCTGAGA
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 3800
TGACCTTTATCGGCAGCTGGCTGGCTTTATATGCGGTCAACTGTTAAAGACCTCGCGACCAGAACTGAACGCGGTCAACAGCCACGCCAAACCGACTCT

AGTGTGCGAGCCAATTCGGGCCAATGATGCGTGTGTTTCATTAGCGGACGATTGACGAGGGCAGAGGCTGCCATGTGCGAATTGGCACGAGTGGACATCGG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 3900
TCACACGCTCGGTTAAGCCCGTTACTACGCACACAAGTAATCGCTGCTAACTGCTCCCGTCTCCGACGGTACACGCTAACCGTGTCTACCTGTAGCC

CACTGTGCGCGTCATCGGAGTCGCGCTGCGGCACCGGTGCGGACCGAGCGGATCCCGGAGTTTGGGCCGAACGACCGGTGTGACACCGCCGGCGGAC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 4000
GTGACAGCGGCAGTAGCCTCAGCGCGGACGCGCTGGCCACGCTGGCTCGGCTAGGGCCTCAAACCCGGCTTGACTGGCCACACTGTGGCGGCCGCTG
T V A V I G V A P A A P V R T E P D P G V W A E L T G V T P P A D

CTGACCGCCGGTGGAGTCGGCTGGACGGCTGTGGGACGCGGCATCGGTGACGACCGGGCGGCATCGCGTCTTCTGGCCCGACGGCGCCG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 4100
GACTGGCGGCCACCTCAGCCGCACCTGCCGGACACCCCTGCGCCGTAGCCACTGCTGGCCCGCCGTAGCCGAGAAGGACCGCGGCTGCCCGGC
L T A A V E S A W T A C G D A G I G D D R A G I G V F L A A D G A G

CGAAGGCAGCGCGGACGAGTTAGCGCGGTGCTCGGCCTCCTCGGCCGTGCGGACCGCGCACGACCCGCTCGCCGAGGCCCTCGCCGCGGTGCGTTC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 4200
CGCTTCGCTGCGCTGCTCAATCGCGCCACGAGCCGAGGAGCCGGGCACGGCCTGGCGGTGCTGGGCGAGCGGCTCCGGGAGCGGCCACGCAAG
E G D A D E L A R V L G L L G P C R T A H D P L A E A L A A V R S

CGGCGAATGCGGCCTGGCCCTCGTCTGGTTCGGACGGTGCCACGCCGTGCTCGCGAGGGGCGACGACCCCGACCGCGACTTCGCCTACCGGCTCGGC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 4300
GCCGCTTACCGCGACCGGGAGCAGCCAGCCTGCCACGGGTGCGGCACGAGCGCTCCCCGCTGCTGCGGGCCTGGCGTGAAGCGGATGGCCGAGCCG
G E C G L A L V G S D G A H A V L A R G D D A R T A T F A Y R L G

ACCGACCGGAGTTGCCTGACAACGTTCTCACCTCGTTCGAGTCCGCCCTGCGCGCGGGCCGCCCGCCCGCGCGCGCGCGCGGAAAGCCCGCCG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 4400
TGGCTGCGGCTCAACGACTGTTGCAAGAGTGGGAGCAGCTCAGGCGGGACGCGCGCCGGCGGGGGCGGGCGCGCGCGCGGCTTCCGGGGCGGC
T D A E L P D N V L T L V E S A L R A G R P A P G R A A A E G P A V

TCCCGATCGCGCTGACGGCGCACAGCCAGGACGCCCTGCGGGCCCGCGCCGCGACCTGTGCATACTCCTCGACCGCGAGCCGGAACCTCCGCTCGCCGA
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 4500
AGGGCTAGCGCGACTGCCCGTGTGGTCTCGGGACGCCGGCGCGCGGCTGGACACGTATGAGGAGCTGCGGCTCGGCCTTGAGGGCGAGCGGCT
P I A L T A H S Q D A L R A R A A D L C I L L D A E P E L P L A D

Anexo I

CGTCGCCTTGTCCCTCGCCACGACCGGCCCGCGCTCGCCGAGCGCCGTGTTCTCGGCGCCGACCGGGCCCGCTCACCGACGAGCTGGCCGCCCTC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 4600
GCAGCGGAACAGGGAGCGGTGCTGGCCGGGGCGGAGCGGCTCGCGCGGCACAAGCAGCCCGGGCTGGCCGGGGCGGAGTGGCTGCTCGACCGCGGAAG
V A L S L A T T G P A L A E R A V F V G A D R A R L T D E L A A F

AGCAAGGGCGGATTACCCGTACCGGGACTCGCCACCGGGCGCCCGCCCGGCCGCGGGCGGGCTCGCCTTCGTGTTCCCGGGCACGGGCCCAAT
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 4700
TCGTTCCCGCCTAATGGGCATGGCCCTGAGCGGTGGCCGCGGGGGCGGGCCGCGCCCGCCGAGCGGAAGCACAAGGGGCCCGTGGCCGGGGTTA
S K G G L P V P G L A T G G P A P A A G G L A F V F P G H G P Q W

GGCCGGGCATGGCCGACGAGTTGATGGACCGCTTCCCGGTGTTGCGCGAGCAGCTGCAGGCCCTGCGCGGACGCCCTCACACCGCACGTGGACTGGCCCGT
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 4800
CCGGCCCGTACCGGCTGCTCAACTACCTGGCGAAGGGCCACAAGCGGCTCGTCGACGTCGCGGACGCGCCTGCGGGAGTGTGGCGTGCACCTGACCGGGCA
P G M A D E L M D R F P V F A E Q L Q A C A D A L T P H V D W P V

CATGGACGTCCTGCGCGGCCGACGGACACGCCGTGCTCGACCGGCCGACGTGGGACAGCCCGCACTCTGGGGATGATGGTGTCCCTGGCCGCCCTG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 4900
GTACCTGCAGGACGCGCCCGGCTGCTGTGCGGCAGCGAGCTGGCCGGCCTGCACCCTGTGCGGGCTGAGACCCGCTACTACCACAGGGACCGCGGGAC
M D V L R G R T D T P S L D R P D V G Q P A L W A M M V S L A A L

TGGCGGTCTGGGGCGTGGTGCCCGACCGCTGATCGGCTCCAGCGCCGGCGAGCTC**CCGGCCGACGGTTCGCGGGCGCACTGTCCGTGAGGACGGCG**
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 5000
ACCGCCAGGACCCCGACACGGGCTGCGGCACTAGCCGAGGTGCGGCGCCTCGAGGGCCGCGCTGCCAGCGCCCGCTGACAGGCAGCTCCTGCCGC
W R S W G V V P D A V I **G S S** A G E L P A A T V A G A L S V E D G A
ATp

CGCGGGCCGTGGCAGTGTTCGGCCAGGGCGAGGTGCACCTGCTCGAACACGGCGTATGCTGTCCGTCTCGCGTCCGCCGACGAGGTACCGGAACGGCT
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 5100
GCGCCCGCACCGTCAAGCCGGTCCCGCTCCACGTGGACGAGCTTGTGCCCACTACGACAGGCAGGAGCGCAGGCGGCTGCTCCATGCGCTTGGCGA
R A V A V F G Q A Q V H L L E H G V M L S V L A S A D E V R E R L

GACCCGCTACGACGGCCTCGACCTCGCCCGGTCACGCCCGGGCTCGGTCTGGTCTCCGGCTCGAGGACCAGGCCGACGAACCTCTCGCCGAACTG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 5200
CTGGCGATGCTGCCGAGCTGGAGCGGGCGGACGTCGGGGGGCCGAGCCAGGACCAGAGGCCGAGCTCCTGGTCCGGCTGCTTGGAGAGCGGCTTGGAC
T R Y D G L D L A A V H A P G S V L V S G V E D Q A D E L L A E L

GCCGTGGACGGCATCAAGGCCAGGGCCATCGACATCCGCATCGCGGCGCACGGCCCGCAGGTTCGACGCGGTCTCGACGGGCTGCGCGCCGACCTCGCCC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 5300
CGGCACCTGCCGTAGTTCGGTCCCGGTAGCTGTAGGCTAGCGCCGCTGCCGGCGGTCCAGCTGCGCCAGGAGCTGCCGACGCGGGCTGGAGCGGG
A V D G I K A R A I D I R I A A H G R Q V D A V L D G L R A D L A P

CGATCGAGCCCCCGCCGACGGACGTGCCGATCTACCTGGGCACCACGGGCGAGCGGGTTCGACCCGCTCACGCTCACCGCGGACCACTGGTGCCGCAACT
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 5400
GCTAGCTCGGGGCGGCTGCCTGCACGGCTAGATGGACCCGTGGTGCCCGCTCGCCAGCTGGGCGAGTGCAGTGGCGCCTGGTGACCACGGCGTTGGA
I E P R P T D V P I Y L G T T G E R V D P L T L T A D H W C R N L

GCGCCACACCGCCGACTTCGAGAGCGCCACCGCGCCGCACTGGCCCGGGCCACCGCTGCTCGTTCGAGTCGAGCCCGCACCCCGTCTCACAGGGCC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 5500
CGCGGTGTGGCGGCTGAAGCTCTCGCGGTGGCGCGGCGTACCAGCGCCCGTGGCGGACGAGCAGCTCAGCTCGGGCGTGGGGCAGGAGTGGTCCCGG
R H T A D F E S A T R A A L A A G H R L L V E S S P H P V L T R A

GTGCAGGAGACCGCCGAGGACCGCCGCTGGACGCCGTGCTCCTCGGGACGCTGCGCCGGGGCCCGGGCAGCGCCGACCACTGACCGCTGCGCTGGCCG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 5600
CACGTCTCTGGCGGCTCCTGCGGCGGCACCTGCGGCAGCAGGAGCCCTGCGACGCGGCCCGGGCCCGTTCGCGGCTGGTGGACTGGCGACGCGACCGGC
V Q E T A E D A A V D A V V L G T L R R G P G S A D H L T A A L A V

TGCTGCACGCGCACGGCGTACTCCCCGACTGGGACGCGGTGTTTCGGGGCCGGGACGCCCCACGGTGCCGCTGCCCGCTATCCGCTGGACGCGCGCC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 5700
ACGACGTGCGCGTGCCGCATGAGGGGCTGACCTGCGCCACAAGCGCCCGCCCTGCGGGCGTGCCACGGCGACGGGCGGATAGGCGACCTGCGCCGCGG
L H A H G V L P D W D A V F A G R D A R T V P L P A Y P L D A A P

CGCAGCCACCCCGCCCGTCCCTCGCCGACGGCTCGCGGCCGCCCGACCGCCACGAACCTCCTCCTCGACCTGGTCCGCGCGCAGGTCCGCGCCGTC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 5800
GCGTCCGGTGGGGCGGGGACGGGAGCGGCTGCGCGAGCGCCGGGCGGGGCTGGCGGTGCTTGAGGAGGAGCTGGACCAGGCGCGCTCCAGCGCGGCGAG
A A T P A P S L A D R L A A A P D R H E L L L D L V R A Q V A A V

CGGGGCGGGGACGCGCGCTCCTCGCCGACGTGGACGCGCCGCTCCTGGAGCTGGGCGTGCAGCTCGGCGGGCGGGTGTCCCTGGCAACCTCCTGAACG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 5900
CGCCCGCCCTGCGCGCGAGAGGCGGCTGCACCTGCGCGGCGAGGACCTCGACCCGAGCTGAGCCGCCCGCCACAGGGACGCGTGGAGGACTTGC
A G G D A A L S A D V D A P L L E L G V D S A G A V S L R N L L N A
ACP

CCGCCACCGCCCTGAAGCTGCCCGTGTCCGCGGTGTTTCGACAACCCGACGTGCCGAGGCGCTGCCACGCACATCGACCGGGTCTGTCTCGGCGCGCCAC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 6000
GGCGGTGGCCGGACTTCGACGGGCACAGGCGCCACAAGCTGTTGGGCTGCACGGCTCCGGAGCGGTGCGTGTAGCTGGCCAGGACGAGCCGCGCGGGTG
A T G L K L P V S A V F D N P T C R G L A T H I D R V L L G A P T

CGACGACGACGAGGCGCGGTCCCATGACGGGCGGCGACGACGAGCGGATCGCCGTCGTCGCGATGAGCTGCCGCTGCCCGCGGGGTACACAGCCCC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 6100
GCTGTGTGTCTCCGCGGCCAGCGTACTGCCCGCGCTGTGCTCGGCTAGCGGCAGCAGCGTACTCGACGGCGGACGGGCGCCCGCATGTGTCCGGG
D D D E A P V A M T G G D D E P I A V V A M S C R L P G G V H S P

GAGGACCTGTGGGAGCTGCTGTCCGCGGGCGGCGACGGCATCACCGGTTCCCGCCGACCGCGGCTGGGACCTGGACGGACGCTACGACGAGGACGCGG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 6200
CTCCTGGACACCCTCGACGACAGGCGCCCGCGCTGCCGTAGTGGCGCAAGGGGCGGCTGGCGCCGACCCTGGACCTGCCTGCGATGTGTCTCTCGGG
E D L W E L L S A G G D G I T A F P A D R G W D L D G R Y D E D A E

AGCGGTGGGCGGCTACTACCAGCGGAAGCCGGATTCTCTGCACGAAGTCACCGAGTTCGACCCGGAGTTCTCCCGGATCTCGCCGCGGAGGCCCTCGC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 6300
TCGCCACCCGCCGATGATGGTCCGCTTCGGCCTAAGGACGTGCTTCAGTGGCTCAAGCTGGGCTCAAGAGGGCTAGAGCGGCGGCTCCGGGAGCG
R V G G Y Y Q R E A G F L H E V T E F D P E F S R I S P R E A L A

CATGGACCCGACGAGCGGCTGTCTCTGGAGATCTCTGGGAGGCGATCGAGCGCGGGGATGGACCCCGGAGGCTGGGGGCGAGCCGCGGGGCGTG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 6400
GTACCTGGGCGTGTGCGCGACGAGGACCTTAGAGGACCTCCGCTAGCTCGCGCGCCCGTACCTGGGGCGTCCGACGCCCCGTCGGCGCGCCCGCAC
M D P Q Q R L L L E I S W E A I E R A G M D P A R L R G S R A G V

TTCGTGGGCGGATGACGATGGACTACGGGCGGACTGCACGAGGCGCCCGGACCTGGAGGGCTACCTCCTCACCGCAACACCGGAGCGTGCCT
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 6500
AAGCACCCGCGTACTGCTACCTGATGCCCGGCGTGCAGTGTCCGCGGGCGCCTGGACCTCCCGATGGAGGAGTGGCCGTTGTGGCGCTCGCAGCGGA
F V G A M T M D Y G P R L H E A P A D L E G Y L L T G N T A S V A S

CCGGCAGGCTCTCTACAGTTCGGCTTCTCAGGGCCGCGCTCACCGTCGACACGGCGTGTGTCGTCGTCGTCGGCCGCCCTGCACCTGGCCGTGCAGTC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 6600
GGCCGTCCGAGAGGATGTGCAAGCCGAAGAGTCCCGGGCGGAGTGGCAGCTGTGCCGACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG
G R L S Y T F G F S G P A V T V D T A C S S S L A A L H L A V Q S
KS

GCTGCGCCGCGGCGAGTGCACCGTTCGCGCTCGCGGGCGGCTCACGGTCATGCCGTCGCGAGGCATGTTCTGTGGAGTTCAGCAGGCGAGCGCGCTGGCC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 6700
CGACGCGGCGCGCTCACGTGGCAGCGGAGCGCCCGCGCAGTCCAGTACGGCAGCGTCCCGTACAAGCACCTCAAGTCGTCGCGCGGCGACCGG
L R R G E C T V A L A G G V T V M P S Q G M F V E F S R Q R A L A

Anexo I

CCCGACGGCCGCAGCAAGGCGTTCTCGGCCAGCGCGACGGGTTTCGGCCTCGCCGAGGGCGCCACCATGGTGTCTCCTCGAGCGGCTCTCCGACGCCCGCC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 6800
GGGCTGCCGGCGTCGTTCGCAAGAGCCGGTTCGCGCCTGCCAAGCCGGAGCGGCTCCCGCGGTGGTACCACGAGGAGCTCGCCGAGAGGCTGCGGGCGG
P D G R S K A F S A S A D G F G L A E G A T M V L L E R L S D A R R

GCCACGGCCACCCCGTGTCTCGCACTGGTCAAGGGCACCGCCATCAACCAGGACGGCGCTCCAACGGCCTCACCGCCCCAGCGGCCCGCGCAGCGCGG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 6900
CGGTGCCGGTGGGGCACGAGCGGTGACCAGTTCCTGGCGGTAGTTGGTCTGCGCGCAGGTTGCCGGAGTGGCGGGGTTCGCCGGGCGCGTTCGCGGC
H G H P V L A L V K G T A I N Q D G A S N G L T A P S G P A Q R R

GGTGTCCGGCAGGCGCTCGCCGACGCGGGCGTCTCCGCGCCGAGGTCGACGCCGTCGAGGCGCACGGCACCGGCACGAAGCTGGGGCAGCCGATCGAG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 7000
CCACTAGGCCGTCGCGAGCGGCTGCGCCCGCAGAGCGGGCGGCTCCAGCTGCGGCAGCTCCCGTGCCTGGCCGTCCTTCGACCCGCTGGGCTAGCTC
V I R Q A L A D A G V S A A E V D A V E A H G T G T K L G D P I E

GCGGACGGCTGTCTGCCACCTACGGCCAGGACACCCGGCCGACCGGCCGCTGTCTCGGCTCGGTGAAGTGAACATCGGCCACACGCAGGCCGCGT
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 7100
CGCCTGCGCGACGAGCGGTGGATGCCGGTCTCGTGGGCGGCTGGCCGGCGACGACGAGCCGAGCCACTTCAGCTTGTAGCCGGTGTGCGTCCGGCGCA
A D A L L A T Y G Q E H P A D R P L L L G S V K S N I G H T Q A A S

CCGGCTCGCGGGCCTCATCAAGATGGTGTCTCGCGTGCAGCACGGCGAGCTGCCAAGAGCCTGCACATCACCGAGCCACCCCGCACGTTGGACTGGAC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 7200
GGCCGACGCGCCCGGAGTAGTCTTACCAGAGCGGACGTCGTGCCGCTCGACGGTTCCTCGGACGTGTAGTGGCTCGGGTGGGGCGTGCACCTGACCTG
G V A G L I K M V L A L Q H G E L P K S L H I T E P T P H V D W T

GACGGCGCGGTGTCTCGTCTCACCGAGACCACCGCTGGCCGGAGGTCGAGCGCCCCGGCGGGTGGTGTCTCCTCGTTCGGCATCAGCGGCACCAAC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 7300
CTGCCCGCCACAGCGAGCAGTGGTCTGTGGTGCAGCACCGGCTCCAGCTGCGGGGGCCGCCAGCCACACAGGAGCAAGCCGTAGTCGCCGTGGTTG
T G A V S L V T E T T P W P E V E R P R R V G V S S F G I S G T N

GCCCACGCCGCTCTGGAGCAGGCGCCGACGAGGTTCCGGCCGACCCGGAGCCGTCGTTCGGCTCCGTGCCGCTGGTGTCTCCGCGCGCTCGGAGGACG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 7400
CGGGTGCAGGACCTCGTCCGCGGCTGCTCCAAGGCCGCTGGCCTCGGCAGCAGCCGAGGCACGGCGACCACAAGAGCGCGCGAGCCTCCTGC
A H A V L E Q A P D E V P A D P E P S S A S V P L V F S A R S E D A

CCCTGCGCGCCACGCGCCGACTCGCCGCTGTCTGGAGCAGTTGACGCCCGCCGACATGGCGGCCGCTGTCTGTCCCGCTCGGTCTTCGAACACCG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 7500
GGGACGCGCGGTGCGCGGGGCTGAGCGGCGACAGCCTCGTCAACTGCGGGCGGCTGTACC CGCGCCAGACGACAGGGCGAGCCAGAAGCTTGTGGC
L R A H A A R L A A V V E Q L T P A D M A A G L L S R S V F E H R

GGCCGTTGGTTCGCGGGCGCGCCGACGCGCTGACGGCGGTGGCCAGGGGGCCGAGTCGGCGCGGTGGCGATGGGTGTTCGCGCGGCTCGCGGCAA
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 7600
CCGGCACCACAGCGCCCGCGGCTGCGGACTGCGCCACCGGTTCCCGCGGCTCAGCCGCGCCACCGCTACCCACAGCGCCCGCGGACGCCCGTT
A V V V A G G A D A L T A V A Q G G R V G A V A M G V A A A C G Q

GGTGGCGTTCGTGTTCCCGGTCAGGGTTCTCAGTGGCTGGGTATGGGTGTGGAGTTGGCGGAGTCGTTCGGGTTCCGTGCCTCGTTCGAGGCGGTGTG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 7700
CCACCGCAAGCACAAGGGCCAGTCCCAAGAGTACCAGCCATACCCACACCTCAACCGCCTCAGCAGCGGCCACAAGGCACGGAGCAACCTCCGCACAC
G G V R V P G Q G S Q W L G M G V E L A E S S P V F R A S L E A C A

CGGAGGCGTTGGCGCGTATGTCTGACTGGTCTGTGGGATGTCTGGCTGACCGCGCCGAGTCGGAGCGGGTTCGATGTGGTGCAGCCCGCGCTGTGGGC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 7800
GCCTCCGCAACCGCGCATAACAGCTGACCAGCAACCCCTACACGACCGACTGCGCGGCTCGACCTCGCCAGCTACACCAGTTCGGGCGCGACCCCG
E A L A P Y V D W S L W D V L A D A A E L E R V D V V Q P A L W A

Anexo I

GGTGATGGTGTGCGCTTGC GCGCTTGTGGCGCTCGTACGGTGTGAGCCGGATGCGGTGGTGGGGCACAGTCAGGGTGAGATCGCTGCGGCGTATGTCGCG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 7900
CCACTACCACAGCGAACGCCGCAACACCGCGAGCATGCCACAGCTCGGCCACCCCGTGTAGTCCACTCTAGCGACGCCGATACAGCGC
V M V S L A A L W R S Y G V E P D A V V **G H S** Q G E I A A A Y V A
ATp

GGTGGCGTGTGCGCTTGC GCGACGCTGCCAAGATCGTGGCGCAGCGGAGCAAGGCCATTCTTGC GCGTGTGCGGGTCAAGGGCGGCATGGTGTGCGCTCGCGGCGA
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 8000
CCACGGCAGCAGCGAACGCCTGCGACGGTTCAGCACCGCGTGCCTCGTTCGGTAAGAACCGCAGCAGCCAGTCCCCTGACACAGCGAGCGCGCT
G A L S L A D A A K I V A Q R S K A I L A L S G Q G G M V S L A A T

CGGTGACGACGCTCGTGAGTGGATCGCGCCGTGGGGCGAGCGGGTCTCGGTGCGCGGGTCAACGGGCGCTGGTTCGGTGGTGGTGGCGGGTACGCGGA
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 8100
GCCAGTGTGCTGCGAGCACTACCTAGCGCGCACCCCGCTCGCCAGAGCCAAACGGCGCCAGTTGCCCGGACCAAGCCACCACCACCGCCACTGCGCCT
V D D A R E W I A P W G E R V S V A A V N G P G S V V V A G D A D

CGCGCTGGACGAGTTGGTGGAGGGCTGTGCCCGCATGGCCGGCGGGCCCGCCGGATCCCGTGC GACTACGCGTGC GACTCGCCGCATGTGCGAGCGCATC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 8200
GCGCGACTGTCAACCACCTCCGCACACGGCGGCTACCGCGCCCGGGCGGCGCTAGGGGCGAGTGTGCGCAGCGTGTGAGCGCGGTACAGCTCGCGTAG
A L D E L V E A C A A D G R R A R R I P V D Y A S H S P H V E R I

GAGGGCGAGTTGGCGAGTGTGCTGGCCGGGGTGGAGCCCGGGTAGCGAGTGTGCCGATGCTGTCGACCGTCAAGGGGAGTGGCTGACCGGTACCGAGG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 8300
CTCCCGCTCAACCCTCACACGACCGGCCACCTCGGGGCCATCGCTCACACGGCTACGACAGCTGGCAGTGGCCCTCACCGACTGGCCATGGCTCC
E G E L A S V L A G V E P R V A S V P M L S T V T G E W L T G T E V

TGGGTGCGGATTACTGGTATCAGAATCTGCGGCACACCGTGC GTTTCGAGGAGGCCACGCGCGGTCTCCTCGACACGGCATCGGCGCCTTCGTGGAATG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 8400
ACCCACGCCTAATGACCATAGTCTTAGACCGCGTGTGGCAGCAAAGCTCCTCCGGTGC GCGCCAGAGGAGCTGGTGGCGTAGCCGCGGAAGCCGCTTAC
G A D Y W Y Q N L R H T V R F E E A T R G L L D H G I G A F V E C

CAGCCCCACCCCGTCTCACCATCGCGGTGCGGAGACCGTGGACCGGACCGGAGGCGACGCCATCGTGTGCGCACCTGCGCCGACCGACGCGGAC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 8500
GTGGGGGTGGGGCACGAGTGGTAGCCGACGCGCTGTGGCACCTGCGCTGGCTCCGCTGCGGTAGCAGCAGCCGTGGGACGCGCGGTGGCTGCGCGT
S P H P V L T I G V R E T V D A T G G D A I V V G T L R R T D G D

CTCGACCGGTTCCTGCTCTCCGCCCGGAGGCGTACGTGGAGGGCTGTCTTCGACTGGCATGCCGTGCGTGCCCGCGGGCGGCACGTGGACTGCCGA
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 8600
GAGCTGGCCAAGGACGAGAGGCGCGGCTCCGCATGCACCTCCCGACAGGAAGCTGACCGTACGGCAGCACGGGCGGCCCGCTGCACCTGGACGGCT
L D R F L L S A A E A Y V E G L S F D W H A V V P A G R H V D L P T

CGTACCCGTTCCAGCGCAGGCGGTTCTGGCTCGACGAGACATCCCGCCCCGCGGGCGGCGACGACCCCGGTTTCGCGGAGTTCTGGGCCGCGGTGA
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 8700
GCATGGCAAGGTGCGTCCGCCAAGACCGAGCTGCTCTGTAGGGCGGGCGCCCGCCCGCTGCTGGGGCGCAAGCGCCTCAAGACCCGCGCCACCT
Y P F Q R R R F W L D E T S R P A A G G D D P A F A E F W A A V E

CGCGGGGACCTCGCCACCACGCTCGACGTGCCCGGACGCGCCCTGACCGACGTCCTGCCAGCCTCGCCGACTGGCGGGCGCCGAGCCAGGAGCGG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 8800
CGCGCGCTGGAGCGGTGGTGGAGCTGCACGGGCGGCTGCGCGGGGACTGGCTGCAGGACGGGTGCGAGCGGCTGACCGCCGCGCGCTGGTCTCGCC
R G D L A T T L D V P A D A P L T D V L P S L A D W R R R S Q E R

ACCGCGTGC GACTCGTGGCGCCACCGGATCACCTGGCGTCCGCTCGGCACCGCCACCACCCACCCCTGACCGGCACCTGGCTCTGGTGGTCCGCGG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 8900
TGGCGGCGAGCTGAGCACCGCGGTGGCTAGTGGACCGCAGGCGAGCCGTGGCGGTGGTGGGGTGGGACTGGCCGTGGACCGAGGACCAGCCAGGCGGCC
T A V D S W R H R I T W R P L G T A T T P T L T G T W L L V G P P A

Anexo I

CCGAGGCGGAGCACCCGTGGCAGACGGCGGCCGCCGACGCGCTGCGTGCCACGGCGCCGACGTCTGCACGTGAAGGATCGGCCGAGTCCGTGAGCGC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 9000
GGCTCCGCCTCGTGGGCACCGTCTGCCGCCGGCGGCTGCGCGACGCACGGGTGCCCGGCTGCAGGACGTGCAGCTTCTAGCCGGCTCAGCGACTCGCG
E A E H P W Q T A A A D A L R A H G A D V L H V E G S A E S L S A

CCTGGACGTCTACCCAGTGGCGTGTCCCTGCTGGCCCTCGACGAACAGCCCGACCCCGGGGCGCCGACCGTCCCTGCGGCCTCTCCGGCACGCTG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 9100
GGACCTGCAGCATGGGTACCCGACGACAGGGACGACCGGGAGCTGCTTGTGGGCTGGGGCCCCGCGGCTGGCAGGGGACGCCGAGAGGCCGTGCGAC
L D V V P S G V L S L L A L D E Q P D P G A P T V P C G L S G T L

GCACCTGGTGGGGCGCTCGGCAGCGCCGGATCGGCGCGCCGCTGTGGCTGGCCACGACGGGCGGGTGTCCGTCCGACGCTCCGACCCGCTGACCCACC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 9200
CGTGACCACGCCCGAGCCGCTGCGGCCCTAGCCGCGCGGACACCGACCGGTGCTGCCCGCGCCACAGGACGCTGCGAGGCTGGGCGACTGGGTGG
A L V R A L G D A G I G A P L W L A T T G A V S V G R S D P L T H P

CGGCGCAGAGCCAGGTGTGGGGCCTCGGCCTCACCGTCCGCTGGAGACCCGGAGCGCTGGGGGGCCCTCGTCGACCTCCCGTGACCCCGACGCCAG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 9300
GCCCGTCTCGGTCCACACCCCGGAGCCGAGTGGCAGCCGACCTCGTGGGCCTCGCGACCCCGCGGAGCAGCTGGAGGGGCACTGGGGGTGCGGTC
A Q S Q V W G L G L T V G L E H P E R W G G L V D L P V T P D A R

GACCGCACCCGGCTCGCCGCGCTCCTCGCCGGCAGACGGACGAGGACCAACTGGCCATCCGTGCCCTCCGGCGTGACGGTGGCCGGCTCGTCCCGGCC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 9400
CTGGCCGTGGGCCGAGCGCGGCAGGAGCGGCCCTGCTGCCTGCTCCTGGTTGACCGGTAGGCACGGAGGCCGCACTGCCACGCGCCGAGCAGGGCCGG
T G T R L A A V L A G Q T D E D Q L A I R A S G V T V R R L V P A

CCCGCACCCGGCGCGCACCCGCCTGGCGTACCAGCGGCACCCGCGTGATCACCGCGGCACCCGGCGCCATCGGCGGGCACGTGCGCCGCTGGCTGGCGG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 9500
GGGCGTGGCCGCCCGCTGGGCGGACCCGATGGTCGCCGTGGCGGACTAGTGGCCGCCGTGGCCCGGTAGCCGCCGCTGCAGCGGGCGACCGACCGCC
P A P A A A P A W R T S G T A L I T G **G T G A I G G H V A R W L A A**
KR

CCAGCGGCGCGAACACATCGTGTGCTGACCAGCCGACGCGGCCCGGACGCCCCGGCGCCAGGAACGGTCCGCGAACCTGACCGCGCTCGGCGCCCGGAC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 9600
GGTCGCCCGGCTTGTGTAGCAGACTGGTGGCGTCCCGGGCCTGCGGGGCGCGGGTCTTGACCAGCGGCTTACTGGCGGAGCCCGGGCCTG
S G A E H I V L T S R S G P D A P G A Q E L V A E L T A L G A R T

CACCGTCCCGCTGCGACGCGCGGACCGCACCGCGCTCGCCGACCTGTTCGCCGGCTCGACGCCGACGGCACCCCGTCCGCTCCGTCTTCCACGCG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 9700
GTGGCAGCGCGGACGCTGCGCGGCTGGCGTGGCGGAGCGGCTGGACAAGCGGGCCGAGCTGCGGCTGCCGTGGGGGACGCGAGGAGGAGGAGTGGCG
T V A A C D A A D R T A L A D L F A R L D A D G T P V R S V F H A

GCGGGCACCGTCCCGTCCCTGCTCCTCGCCGACACCGCCACCTCGGACCTCGCTACGCCGTGGCGCGAAGGCCGTGGCGCGGTCCACCTCGACGAGC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 9800
CGCCCGTGGCAGGGCAGGGACGAGGAGCGGCTGTGGCGGTGGAGCCTGGAGCGGATGCGGCAGCCCGGCTTCCGGCAGCCCGCCAGGTGGAGTGTCTG
A G T V P S L L L A D T A T S D L A Y A V G A K A V G A V H L D E L

TGTGCGCCGACGCGAGTTGGACGCGTTCTGCTGTTCTCCTCCGGCTCCGCGGTGTGGGGCAGCGCGAGCTGGGGGCGTACGGCGCCGCAACGCCTT
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 9900
ACACGCGCCCTGCGCTCAACCTGCGCAAGCAGACAGAGAGGCGGAGGCGCCACACCCCGTCCCGCTCGACCCCGCATGCCGCGCGGTGCGGAA
C A G R E L D A F V L F S S G S A V W G S G E L G A Y G A A N A F

CCTCGACGGCTCGCGCAGCGCCACCGCCGACGGACTGCCCGCCACCTCGGTCTCCTGGGGCATGTGGGGGGCGGGCAGTGGTACCGCGCACCTG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 10000
GGAGCTGCCGAGCGCTCGCGGTGGTGGGCTGCCTGACGGGCGGTGGAGCCAGAGGACCCCGTACACCCCGCCCGCCGCTACCAGTGGCCGCTGGAC
L D G L A Q R H H A D G L P A T S V S W G M W A G G G M V T G D L

Anexo I

SacI

GACGACCAGCTGCGCCGCGGGGCTGCGCCCCATGCGGCCGAGCTCGCGGTCGGCGCGCTCGCCACCGCCCTCGCGGGGGGCGAGACCACCTCACCG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 10100
CTGCTGGTTCGACGCGGGCCCGGACGCGGGGTACGCCGGCTCGAGCGCCAGCCGCGGAGCGGTGGCGGGAGCGCCCGCCGCTCTGGTGGGAGTGGC
D D Q L R R R G L R P M R P E L A V G A L A T A L A A G E T T L T V

TCGCGGACTTCGACTGGGCCCGCTTCGCCCCCGGCTTCACGGCCGCCGCTCCCGCCCTGATCAGCGAGATACCCGACGTCGTCGCGCTCGCCGACA
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 10200
AGCGCCTGAAGCTGACCCGGGCGAAGCGGGGCGAAGTGC CGGCGGGCGAGGGCGGGGACTAGTCGCTCTATGGGCTGCAGCAGCGGAGCGGGCGTGT
A D F D W A R F A P G F T A A R S R P L I S E I P D V V A L A A Q

GGACGCGGTGGACCAGCCCGCCCGGTGGCGCGCTCGCCGACCGCCTCGCCGGGCTCACCGACGCCGACCGGGCGCGGCTCCTGCTCGACTGGTCCGC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 10300
CCTGCGCCACCTGGTGGGCGGGCGCCACCGCGCGAGCGGCTGGCGGAGCGCCCGAGTGGCTGCGGCTGGCCGCGGCGGAGGACGAGCTGGACCAGGCG
D A V D Q P A A G G A L A D R L A G L T D A D R R R L L L D L V R

GGCCACGCCCGCGGTGCTCGGCCACGAAGACCCGACACCGTCACGCCGACCGCGGTTCCGCGAGCTGGGCTTCGACTCCCTGACCGCCATCGAAC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 10400
CCGTTGCGGGCGCCACGAGCCGGTGTCTCCTGGGCTGTGGCAGTGGCGGCTGGCGCGCAAGGCGCTCGACCCGAAGCTGAGGGACTGGCGGTAGCTTG
G H A A A V L G H E G P D T V T P D R A F R E L G F D S L T A I E L
ACP

TCCGCAAGCGCCTCAACACCGGACCGGCGTACGACTGCCCGGACCGTCTGTTGACCACCGACCCACGGCCCTCGCCGCGAACTGCACACCGT
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 10500
AGGCGTTCCGCGAGTTGTGGCGCTGGCCGCATGCTGACGGGCGCTGGCAGCACAAGCTGGTGGGCTGGGGGTGCCGGGAGCGGCGGCTTACGTGTGGCA
R K R L N T A T G V R L P A T V V F D H P T P T A L A A E L H T V

CGTGTTCGGCGAACAGCAGACCGCCACGCCGACCCCTGGCCACCACGACCCGGGCGCGGACGAAACCGATCGCCGTGGTGGCGATGAGCTGCCGCTACCC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 10600
GCACGAGCCGCTTGTGCTCTGGCGGTGGGCTGGGACCGGTGGTGTGGGCCCGGGCGGCTGCTTGGCTAGCGGCACCACCGCTACTCGACGGCGATGGGG
V L G E Q Q T A T P T L A T T T R A A D E P I A V V A M S C R Y P

GGCGCGTCCGCGGCCCGGAGGACCTGTGGTTCGCTCGTATGGCCGACAGCAGCCATCGGTGCGTTCGCCGACGACCGGGGCTGGGACACCGACCGGC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 10700
CCGCCGAGGCGCGGGGCTCCTGGACACCAGCAGCAGTACCGCCCTGTGCTGCGGTAGCCACGCAAGGGGCTGCTGGCCCCACCTGTGGCTGGCCG
G G V R G P E D L W S L V M A G H D A I G A F P D D R G W D T D R L

TGTACCCGGCCGACGCGGGCGCCCTGGGCCGACCCGACCCCTGGAGGGCGCCTTCTCGACGACGCGGGCGGCTTCGACGCGGGCTTCTTCGGCATCGC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 10800
ACATGGGCCGCTGCGCCCGGGACCCGCGCTCGGCGTGGGACCTCCCGCGAAGGAGCTGCTGCGCCCGCGAAGCTGCGCCCGAAGAAGCCGTAGCG
Y P A D A A A L G R S R T L E G A F L D D A G G F D A G F F G I A

CCCGCTGAGGCCCTCGCGATGGACCCGACGAGCGGCGAGTGTCTCGAATGGCTGGGAGGCCCTTCGAACGGGCGGGCATCGACCCGCTCCTCGTGGC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 10900
GGGCGCACTCCGGGAGCGCTACCTGGGCGTCTGCGCGTCCACGAGCTTGACCGGACCCCTCCGGAAGCTTGCCCGCCGCTAGCTGGGCGAGGACGCGG
P R E A L A M D P Q Q R Q V L E L A W E A F E R A G I D P S S V R

GACAGCCTCACCGGCGTTCGTCGGCGCCTCGCCGACAGCTACTCGGGACCATGGAGCAGGCGACGCCCGAGATGGACGGCTACCGCCTGACCGGCG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 11000
CTGTGGAGTGGCCGACAAGCAGCCGCGGAGCGGCTCTCGATGAGCCCGTGGTACCTCGTCCGCTGCGGGCTTACCTGCCGATGGCGGACTGGCCGC
D S L T G V F V G A S P Q S Y S G T M E Q A T P E M D G Y R L T G D

ACGCCCTGAGCGTCTGTCCGGCCCGCTCGCCTACTCGCTGGGCCTGCGCGCCCCGCCGTACCCTCGACACCGCCTGCTCGTCTCGTGGTCCGCT
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 11100
TGGGGGACTCGCAGCACAGGCCGGCGCAGCGGATGAGCGACCCGGAGCGCCGGGGGGCAGTGCCAGCTGTGGCGGACGAGCAGGAGCGACCAGCGGGA
A L S V V S G R V A Y S L G L R **G P A V T V D T A C S S S L V A L**
KS

GCACCTCGCCGTGCAGGCCTGCGGTCCGGGGAGTGCTCGATGGCCCTGGCGGGCGGCACCGCATCATGGTACGCCCTGCCCTTCGCCGAGTTCAGC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 11200
CGTGGAGCGGCACGTCCGTGACGCCAGCCCCCTCACGAGCTACCGGGACCGCCCGCTGGCGCTAGTACCACTGCGCGACGGAAGCGGCTCAAGTGC
H L A V Q A L R S G E C S M A L A G G T A I M V T P L P F A E F S

CTGCAGGGCGCCTGGCCCCGACGGCCCGTGAAGCCCTTCGCCGACGCCCGACGGCGTCCGGTGGGGCGAGGGCGGGCCTGATCCTGCTGGAAC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 11300
GACGTCCCGCCGACCGGGGGCTGCCGGCCACGTTCCGGGAAGCGGCTGCCGCGGCTGCCGACCGCCCGCTCCCGCGCCCGGACTAGGACGACCTTG
L Q G G L A P D G R C K P F A D A A D G V G W G E G A G L I L L E R

GCCTCTCCGACGCCCAGGCCAACGGCCACGAAGTGCTCGCCGTCTCCGGCTCCGCCATCAACCAGGACGGCGCCTCCAACGGCCTCTCCGCGCCAA
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 11400
CGGAGAGGCTGCGGGTCCGGTTGCCGGTGTTCACGAGCGGCAGCAGGCGCCGAGGCGGTAGTTGGTCTGCGCGGAGGTTGCCGGAGAGGCGGGGTT
L S D A Q A N G H E V L A V V R G S A I N Q D G A S N G L S A P N

CGGGCCCGCACAGCAGCAGGTATCCGCGGGCGCTCGCCAACCGGGCCTGTCCGCGCGGACGTGGACGCCCTCGAGGGCCACGGCACCGGCACAAAG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 11500
GCCCCGGCGTGTCTGTCCAGTAGGCGCGCCGCGAGCGGTGCGCCCCGACAGGCGGCGCCTGCACCTGCGCGAGCTCCCGGTGCCGTGGCCGTGTTTC
G P A Q Q Q V I R A A L A N A G L S A A D V D A V E G H G T G T K

CTCGGCGACCCATCGAGGCCAGGCACTCCTTGCCACCTACGGCCAGGAACGCCCGCCGACCAGCCGCTGCTGCTCGGCGCCTGAAGTCGAACATCG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 11600
GAGCCGCTGGGGTAGTCCGGTCCGTGAGGAACGGTGGATGCCGGTCTTGCGGGCGGCTGGTCCGCGACGACGAGCCGCGGACTTCAGCTTGTAGC
L G D P I E A Q A L L A T Y G Q E R P A D Q P L L L G A L K S N I G

GCCACAGCAGGGCGCCTCGGGCATCGGGCGTCAAGATGGTGCAGGCACTGCGCCACGGCGTACTGCGCGCACCTGCACCTGGACCGCCCGTCC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 11700
CGGTGTGCGTCCGCCGAGCCGTAGCGCCCGCAGTAGTTCACACGTCCGTGACGCGGTGCCGATGACGCGCGTGGGACGTGGACCTGGCGGGCAG
H T Q A A S G I A G V I K M V Q A L R H G V L P R T L H L D R P S

CCAGCACGTGGACTGGTGCACCGGCGCCGTGGAACCTCCGCGACGACCGGCCGTGGCCGAGGCCGACCGCCCCGGGGGGCCCGTGTCCGCGTTC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 11800
GGTCTGACCTGACCAGCTGGCCGCGGCACCTTGAGGAGCGGCTGCTGGCCGGCACCGGGCTCCGGTGGCCGGGGCCCGCCGGCGCACAGGCGCAAG
Q H V D W S T G A V E L L A D D R P W P E A D R P R R A A V S A F

GGCGTCAGCGGCACCAACGGCCACGTATCCTCGAACAGGCCGACGTCCAGGCCGAGTCCGTCCCCGCGGACCCGACCGCCGAGCCCGAGCCGCGGACGC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 11900
CCGCGTCCCGTGGTTGCCGGTGCAGTAGGAGCTTGTCCGGTGCAGGTCCGGCTCAGGCAGGGGCGCCTGGGCTGGCGGCTCGGGCTGCCGGCCTGCG
G V S G T N G H V I L E Q A D V Q A E S V P A D P T A E P D G R T L

TGACGTGGGCGCTGTCCGCCCGGACACCGCCGCTGCGCGACAGGCCGCCCGCTGGCCGCCACCGTGCAGGAGCGGGCCCGCCTGTCCCGCTGGA
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 12000
ACTGCACCCGCGACAGGGGGCGCTGTGGCGCGCAGCGCTGGTCCGGCGGGCGGACCGCGGTGGCAGCTGCTCGCCGGGCCGACAGGGGCGACCT
T W A L S A R D T A A L R D Q A A R L A A T V D E R P G L S P L D

Anexo I

CGTCGCCACGACGCTCACCACCGGGCGCGCCGCGCTCGACCACCGCGCCGTCGTCTCGGCACCGACACCTCCGAACTGCGCGTCCGACTCGACGCGCTG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 12100
GCAGCGGTGCTGCGAGTGGTGGCCCCGCGCGCGAGCTGGTGGCGGGCAGCAGGAGCCGTGGCTGTGGAGGCTTGACGCGCAGCCTGAGCTGCGCGAC
V A T T L T T G R A A L D H R A V V L G T D T S E L R V G L D A L

ATCGTCGGCGCGACGACCCGCGCTGGTGC GCGGCACCCCGACGCGCGGGCTCGCCGTGCTGTCTCCGGCCAGGGCTCCAGCGCATCGGCATGG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 12200
TAGCAGCCCGCGCTGCTGGGCCGCGCACCGCGCCGTGGGGGCTGCCGCGCCCGAGCGGCACGACAAGAGCCGGTCCCGAGGGTCCGCTAGCCGTACC
I V G G D D P A V V R G T P D G G G L A V L F S G Q G S Q R I G M G

GCCGCGGCCTGTACGCGCGCACCCCGTGTTCGCGGACGCCCTCGACGAGGCGTTCGCCACCTCGACCCGCACCTCGCCGTGCCGTGCGCGACGCTGT
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 12300
CGGCGCCGACATGCCCGCGTGGGGCACAAGCGGCTGCGGGAGCTGCTCCGCAAGCGGGTGGAGCTGGGCGTGGAGCGGCACGGCGACGCGTGCAGCA
R G L Y A A H P V F A D A L D E A F A H L D P H L A V P L R D V V

CTTCGGCGACGACAGGCCCTGCTGGACCGCACCGAGTACGCCAGCCCGCGCTGTTCGCCGTGAGGTGCGCCCTTACCGGCTGTGGGAGTCTGGGGT
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 12400
GAAGCCGCTGCTGGTCCGGGACGACCTGGCGTGGCTCATGCGGGTCCGGCGCGACAAGCGGCAGCTCCAGCGGGAGATGGCCGACACCTCAGGACCCCA
F G D D Q A L L D R T E Y A Q P A L F A V E V A L Y R L W E S W G

CTCAAGCCCGACGCGCTCGCGGGCCACTCCATCGGCGAACTCACCGCGCACACGTGCGCCGGGTCTTCTCGTGGCCGACGCTGCGCGATCGTCGCCG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 12500
GAGTTCGGGTGCGCGAGCGCCCGGTAGGCTAGCCGCTTGGTGGCGCGTGTGCAGCGGCCCCAGAAGAGCGACCGGCTGCGGACGCGCTAGCAGCGGC
L K P D A L A **G H S** I G E L T A A H V A G V F S L A D A C A I V A A
ATp

CCCGCGCGCACTCATGCAGGAAGTGGCCCCGGTGGCGTTCATGGTTCGCGGTGCAGGCCCGGAGGACGAGATCGCGCCGGAGCTCCGAGCAGCTCGGCAT****
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 12600
GGGCGCCGCTGAGTACGCTTGTACCGGGGCCACCGCAGTACCAGCGGCAGTCCGGCGGGTCTCTGCTCTAGCGCGGCTCGAGGCTCGTGCAGCCGTA
R G A L M Q E L A P V A S W S P C R P P R T R S R R S S E H V G I
SacI

CGCCGCATCAACGGCCGAACTCCCTGGTCTGTCCGGCAGCGAACCAGAGTGTTCGACCTCGCCGACCGCCTCAAGGCACAGGGCCGCAAGGTGACC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 12700
GCGCGGTAGTTGCCGGGCTTGAGGGACCAGGACAGGCCGCTGCTGGGCTCCACGAGCTGGAGCGGCTGGCGGAGTTCGCTGCCGGGCTTCCACTGG
A A I N G P N S L V L S G S E P E V L D L A D R L K A Q G R K V T

AGGCTGCGCGTACGCCACGCGTTCCACTCGCCGCTCATGGAGCCGATGCTGGCCGCTTCCGCGGTGCTGGCAGGGCCACCGCCATCCGCGGACCC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 12800
TCCGACGCGCAGTCCGTCGCAAGGTGAGCGCGCGTACCTCGGCTACGACCGCGCAAGCGGCACAGCACCCTCCGCGGTGGCGGGTAGCGGGCTGGG
R L R V S H A F H S P L M E P M L A A F R R V V A G A T A H P P T L

TGCCCATCGTGTCCACGCTGACCGCGCGCCGGTTCAGCGCCGAGGAACTGGCCGCGCCGACCACTGGGTGCGGCACGTGCGCCAGCCGGTTCGCG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 12900
ACGGGTAGCACAGGTGCGACTGGCCGCGCGCCAGTCCGCGCTCCTTGACCGCGCGCCCTGGTACCCACGCGTGCACGCGGTCCGCCACGCCAAGCG
P I V S T L T G A P V S A E E L A A P D H W V R H V R Q P V R F A

CGACGGCGTACGGCACTCGTCCGCGAGGGCGTGACGACGATGGTGGAGGTGCGGGCCCGCACCGTGTTCGCCGCGATGGGCCAGGACAGCGCCCCGAC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 13000
GCTGCCGACGTCGCGTGGAGCAGCGCTCCCGCACTGCTGCTACCACCTCCAGCCCGGGCCGTGGCACGAGCGGCGCTACCCGGTCTGTCGCGGGGGCT
D G V T A L V A Q G V T T M V E V G P G T V L A A M G Q D S A P D

Anexo I

ACGGTGTTCGTGCCGACGCTGCGCTCCGACCGGCCGAACCGGACGCGCTCACCTCGCCCGCGCACCTGTACGTCCGCGGCGGAAGCTGGACGCCG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 13100
GTGCCACAAGCACGGCTGCGACGCGAGGCTGGCCGGGCTTGGCCTGCGCGAGTGGGAGCGGCGCGTGGACATGCAGGCGCCGCGCTTCGACCTGCGGC
T V F V P T L R S D R P E P D A L T L A A A H L Y V R G A K L D A A

CGGTGCCCCCTGCGCGCCCGGCCGCGGGTGCACCTGCCACGTACGCTTCCAGCACGACCGCTACTGGCTGACCGCCGCGCTCGCCCCGCGGGCGA
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 13200
GCCAGCGGGAACCGCCGGGCCCGCGGCCAGCTGGACGGGTGCATGCGGAAGGTCGTGCTGGCGATGACCGACTGGCGGCGGCGAGCGGGGGCGCCCGCT
V A L A G P G R R V D L P T Y A F Q H D R Y W L T A A V A P A G D

CGGGGCGGACGCGCCCTCGGCGTCCGCGGGGCGACCACCGCTGCTGCGCGCCGCGGGTCCGCTGCCGACTCCGACGCGCTGCTTTCACCGGCAGG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 13300
GCCCGCCGCTGCGCGGAGCCGACGCGGCCCGCTGGTGGGCGACGACGCGCGGCCAGCGCGACGGGCTGAGGCTGCCGACGAGAAGTGGCCGCTCC
G A A A A L G V A A G D H P L L R A A V A L P D S D G V V F T G R

CTCTCACCCGGACCCATCCGTGGCTCGCCGACCACGGCGTGGCGGCCGGGTGCTGTTCCCGGCACCGGCTTCTCGACCTCGCCCTGCACGCCGCGA
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 13400
GAGAGGTGGGCTGGGTAGGCACCGAGCGGCTGGTGCAGCCGCGGCCACGACAAGGGGCCGTGGCCGAAGGAGCTGGAGCGGGACGTGCGGCGCT
L S T R T H P W L A D H G V G G R V L F P G T G F L D L A L H A A S
DH
PstI
GCCACTGCGGCCTCGACACCGTGCAGAACTGACCCTGCAGGCCCCGATGGCCCTGCCGAGCGCGGGCCCTCGCCCTGCGCGTACCGTGGGCGCGGA
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 13500
CGGTGACGCCGAGCTGTGGCAGCTGCTTACTGGGACGTCCGGGGTACCGGGACGGCTCGCGCCCGGAGCGGGACCGCGAGTCCACCCGCGCCT
H C G L D T V D E L T L Q A P M A L P E R G G L A L R V T V G A E

GAACGACGGGACGCGGAGCATCCGGGTGCACTCACGGGCGGACGACGCCGACCGGGACGCGCCGTGGACCGAGCACGCCACCGGCCTGCTCACACGGGC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 13600
CTTGCTGCCCTGCGCCTCGTAGGCCACGTAGTGCCCGCCTGCTGCGGCTGGCCCTGCGCGCACCTGGCTCGTGGGTGGCCGACGAGTGGTGC
N D G T R S I R V H S R A D D A D R D A P W T E H A T G L L T T G

GACGTGACGAGACCGACCTCGCGGCGTGGCCGCGCGGGGCGGACACCGATGGACGTGACGCGCATCTACGACCGCCTCGACGAACTCGGCTACGGCT
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 13700
CTGACGTCTCTGGTGGAGCGCCGACCGCGGCGCCCGCGCTGTGGTACTCTGACGCTGCCGTAGATGCTGGCGGAGCTGCTTGGCCGATGCCGA
D V Q E T D L A A W P P R G A T P M D V D G I Y D R L D E L G Y G Y

ACGGGCCCGTGTTCGCGCGCTGCGCGGGCCTGGAGCCTCGACCGTGACGTGTACGTGAGGTGGCCCTGCCGACGGGACCGTCCGGCAGCGCTTCAG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 13800
TGCCCGGGCACAAGGCGCGGACGCGCGCCGACCTCGGAGCTGGCACTGCACATGCAGCTCCACCGGACGCGCTGCCCTGGCAGCCGCTGCCGAAGTC
G P V F R A L R A A W S L D R D V Y V E V A L P D G T V G D G F S

CGTCCACCCCGCCCTGCTCGACGCGCGCTGCACGCGCCCGTGTGCGCGCCCTCGACGAGGTGCGGCGCCCGCCCGGCTGCCGTTTCAGCTTCAGGGGA
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 13900
GCAGGTGGGGCGGACGAGCTGCGGCGGACGTGCGGGGACGACGCGGGGAGCTGCTCCAGCCGCGCCGGGGGCGGACGCAAGTCGAAGTCCCT
V H P A L L D A A L H A P V L R A L D E V G A G P R L P F S F R G

GTGCGGCTGCACGCGCTCGGCGCCTCGGCGTGCAGGCTGGTGCAGGCGCGGCGAGGGGCGAGAGATGTCGCTGTCGCTCGCGGACCCGCGC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 14000
CACGCCGAGTGGCGGAGCCGCGGAGCCGCGGACGCCACCGGAGCGGGTCCCGTCCCGTCTTGTCTACAGCGACAGGCAGCGCCTGGGCGGCG
V R L H A V G A S A L R V R W S P T G Q G Q N E M S L S V A D P A G

Anexo I

GGCGGCCGGTGGCGACCGTTCGAGTCACTGGTGATGCGACCGCGGACCTCCGCGGGACCGCTCCCGACCTCCGACGCACTGCTCGAAGTGGAGTGGGC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 14100
CCGCGGCCACCGCTGGCAGTCACTGACCACTACGCTGGCCCGCTGGAGGCGCCCTGGCGGAGGGGCTGGAGGCTGCGTGACGAGCTTACCTCACCCG
R P V A T V E S L V M R P A D L R G T A S P T S D A L L E V E W A

GCCGCTTGCCGCGACCGGGCAGCGCGGGTTCGGGCTCGCGGTGCTCGCGCGGACCGCCTTGGCACGCGACCGCCTCGGGGCGGACCGCCTCGGCCTGGAC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 14200
CGGCGAACGGCGCTGGCCCGTCCCGGCCAGCCCGAGCGCCACGAGCCGCGCTGCCGGAACCGTGCCTGCCGAGCCCCGCTGCCGAGCCGGACCTG
P L A A T G Q R P V G L A V L G A D G L G T D G L G A D G L G L D

GCGGCCGACTTCGGCCTCGACGCCCCACCGTCTCCGACTTCGCGCCCTCGACGCGGTGACCGACCTGCTCGCCCGCTGCCCCGAGCGTCCGACGCGG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 14300
CGCCGGCTGAAGCCGAGCTGCGGGGGTGGCAGAGGCTGAAGCGCGGGAGCTGCCGCACTGGCTGGACGAGCGGGCGACGGGGGCTCGCAGGCTGCGCC
A A D F G L D A P T V S D F A A L D G V T D L L A R C P R A S D A D

ACGGGAAGGGCGCGGGACCGGTGACCTGGGCGCTCGCACTGGCCAGGCTGGCTCGCCGACGAGCGGACCACGAACACGCGCCTGGTGTTCGTGAC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 14400
TGCCCTTCCGCCGCGCCTGTGCCACTGGACCCGCGAGCGTGACCGGGTCCGGACCGAGCGGCTGCTCGCCTGGTGTGTGCGCGGACCACAAGCACTG
G K A A R D T V T W A L A L A Q A W L A D E R T T N T R L V F V T

CCGCGCGCCGCTCGACACCGAGTCCCGACCCGGCGCAGGCGCCGCTGTGGGGCCTGGTCCGGTTCGGCACAGACCGAGAACCCCGCGCGGTTTCGTGCTC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 14500
GGCGCCGCGCGCAGCTGTGGCTGCACGGGCTGGCCCGCTCCGGCGGCACACCCCGACCGAGCCAGCCGCTGTCTGGCTCTTGGGGCCGCGCAAGCACGAG
R G A V D T D V P D P A Q A A V W G L V R S A Q T E N P G R F V L

GTCGACCTCGACGCGCCCGCGAGTCCGCCCTCGCCGTGCCCGCGCTGGCCTCCGACGAGCCCCAACTGGCCGTACGGGCGGGCCGGGTGACCGTGC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 14600
CAGCTGGAGCTGCGGGCGCGGCTCAGGCGGGAGCGGCACGGGCGCGGACCGGAGGCTGCTCGGGGTTGACCGGCATGCCCGCCCGGCCACTGGCAGC
V D L D A A A E S A L A V P A A L A S D E P Q L A V R A G R V T V P

CCAGGCTCGCCCGCACCCCGCGCAGGCTGCTCACCCCGCCCGCGACGCGCCGCTGGCGCCTCGACGTCCAGCCGCGCGGCTCCCTGGACGGTCTCGC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 14700
GGTCCGAGCGGGCGTGGTGGCGGCTGTCGACGAGTGGGGCGGGCGCTGCCGGGCACCGCGAGCTGCAAGTCCGGCGCGCGGAGGGACCTGCCAGAGCG
R L A R T T A D R L L T P P A D G P W R L D V Q P R G S L D G L A

CCTGGTGCACACCCGACGCCACCGCGCCCTCGCACCCGGGAGGTACGCGTCCGCGTCCGCGCGGAGGACTGAACCTCCGCGACGCGCTGATCGCC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 14800
GGACCACGGTGTGGGCTGCGGTGGCGCGGGGAGCGTGGGCGGCTCCATGCGCAGCGGAGCGCGCCGCTCCTGACTTGGAGGCGCTGCGCGACTAGCGG
L V P H P D A T A P L A P G E V R V A V R A A G L N L R D A L I A

CTCGACGCCCCACCCGACGCCCGGAACCGGGCAGCGAGGTCCGCGGCGTGGTCCAGGAGTCCGCTCGGCGGTGCCCGGCTGACCGTGGGCGACCAGC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 14900
GAGCTGCGGGTGGGGTGGCGGCTGGCCCGTCCAGCGCCCGCACCAGTGGCTCCAGCCGAGCCGCCACGGGCGGACTGGCACCCGCTGGTGC
L D A H P D A A E P G S E V A G V V T E V G S A V P G L T V G D H V

TGCTCGGCCTGGTACCAAGGGCTTCGGACCGCTCGCGGTACCGACCACAGGCTCCTCGCCCGGTGCCCGGGGCTGGTCTTCGCGGACCGCGCATC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 15000
ACGAGCCGACCGAGTGGTTCCTGAAGCCTGGCGAGCGCCAGTGGCTGGTGTCCGAGGAGCGGGGCCACGGGCGCCCGACCGAAGCGCTGCGGCGTAG
L G L V T K G F G P L A V T D H R L L A P V P A G W S F A D A A S

GGTGCCTGGCCTTCTTACCGGCTACTACGCCCTGCGCGACCTGGCGGACCTGCGCGCGGAGTCCGCTCCTCGTCCACGGGGCGGGGCGGACGTC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 15100
CCACGGCGACCGAAGGAGTGGCGCATGATGCGGGACCGCTGGACCGCCTGGACCGCGCGCCGCTCAGGAGGAGCAGGTGCCCGCGCCCGCTGCGAG
V P L A F L T A Y Y A L R D L A D L R R G E S V L V H G A A G D V

Anexo I

GGCCTGGCGGCCGTGCAACTCGCCCGGCACTGGGGCGCGGAGGTGTTTCGCGACCGCGCTCCCGAGGAGTGGGACGTCGTACGGGGCGCGGGTGTGCCCG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 15200
CCGACCGCCCGCACGTTGAGCGGGCCGTGACCCCGCGCTCCACAAGCGCTGGCGCGCAGGGCTCCTCACCTGCAGCATGCCCGCCGCCACACGGGC
G L A A V Q L A R H W G A E V F A T A R P E E W D V V R A A G V P D

ATGCGCACCTGGCGTCGTGCGACGACACGGACTTCGCGCGGAAGTTCGCCGTGGAGACCGGGCGGGCGGCGTGCACGTGGTGTGAACCCGCTGGTGGC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 15300
TACGCGTGGACCGCAGCAGCCTGCTGTGCCTGAAGCGCCGCTTCAAGCGGCACCTCTGGCCGCCCCGCGCCGACGTGCACCACGACTTGGGCGACCACGC
A H L A S S D D T D F A A K F A V E T G G R G V H V V L N P L V R

GGAGTTCGCGGGCCCTCCCCCAAGCGCGGAACACGCGGAACCGCCCGGAAGCACGGGAGTTCACCGACGCCTCCCTGGAACGCTGCCCGACGGC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 15400
CCTCAAGCGCCCGGAGGGGGTTCGCGCCCTTGTGCGCCTTGGCGGGGCTTCGTGCCCTCAAGTGGCTGCGGAGGGACCTTGCAGCAGGGCTGCCG
E F A A A S P Q A R E H A E T A P E A R E F T D A S L E L L P D G

GGACGGTTCATCGAGCCGGCAAGGCCGAAGGACACGAGCCGACGCCCTGCCGCCGGCGTCAAGTACGCCGCTTCGACCAGCCGACCGGGCCCGC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 15500
CCTGCCAAGTAGCTCGGCCGTTCCGGCTTCTGTGCTCGGGCTGCGGGACGGCGGGCCGAGTTCATGCGGCAGCTGGTTCGGCGTCCGCCGGGGC
G R F I E P G K A E G H E P D A L P P G V K Y A A F D Q P Q P G P D

ACCGCTGCCAGGAGATGCTCGCCGACGTCCTCGCGCTCTGCGAGGGCGGCGTGTGAACCTCCCTGCCGGTGGCGCCCTGGGACGTCGCCCGCGCACCCGA
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 15600
TGGCGACGGTCTCTACGAGCGGCTGCAGGAGCGCGAGACGCTCCGCCCGCACGACTTGAGGGACGGCCACGCGCGGACCCTGCAGGGCGCGGTGGGCT
R C Q E M L A D V L A L C E A G V L N S L P V R A W D V R R A P E

GGCGTTCGGCACGTCGGCGTGGCCCGGTTACCGGCAAGGCCGTGCTGACCGTGCCCGCACCGATTGACCCGGACCGCCCGGTCCTGGTGACCGGGCGC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 15700
CCGCAAGGCCGTGCAGCCGACCGGGCCAAGTGGCCGTTCCGGCACGACTGGCACGGCGTGGCTAACTGGGCTGGCGGGCCAGGACCCTGCGCCCGC
A F R H V G V A R F T G K A V L T V P A P I D P D R P V L V T G **G**

CTCGGCACGCTCGGGCGGATCGTGGCCCGGACCTCGTCGCGAACGAGGGCTCCGTAACCTCGTCTCGCGGGCCGCAAGGGCCTGGCGACCGAGGGCG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 15800
GAGCCGTGCGAGCCGCTAGCACCGGGCCGTGGAGCAGCGGCTTGTCTCCGAGGCATTGGAGCAGGAGCGCCCGGCTTCCCGGACCGCTGGCTCCCGC
L G T L G A I V A R H L V A E R G V R N L V L A G R K G L A T E G A
KR

CCGAGGAACGCGCGCAACTGGCCCGCTGGGGCGCGAGGTGTCCTGGCGGCCTGCGACGTGGCCGACCGGGACTCGGTGCTGGCGCTCTTCGCCTC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 15900
GGCTCCTTGCAGCGCTGCTTACCGCGCGACCCCGGGCTCCACAGCACCGCCGAGCGTGCACCGGCTGGCCCTGAGCCAGCACCGCGAGAAGCGGAG
E E L R D E L A A L G A E V S V A A C D V A D R D S V V A L F A S

GCTGGACCGGCCGCTCGGGCGGTCGTGCACACCGGGCGTCTGCGACGACGGCGTACTGGAGTCTGCTGACGCCCGACCGGTCGACCGGCTCTTCGGC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 16000
CGACCTGGCCGGCAGCCGCGCCAGCACGTTGGCGCCCGCAGCAGCTGCTGCCGATGACCTCAGCGACTGCGGGCTGGCGCAGCTGGCGCAGAAGCCG
L D R P L G A V V H T A G V V D D G V L E S L T P D R V D R V F G

CCGAAGGTCGACGGCCTGCTGCACCTGGACGAGCTGACCAGGGACCGGACCTGTCCGCGTTCGTGCTGTTCTCCTCGGGCGCGGAGTCTTCGGCAGCG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 16100
GGCTTCCAGCTGCCGACGACGTTGACCTGCTGACTGGTCCCTGGCGCTGGACAGGCGCAAGCAGCACAAGAGGAGCCCGGCTCAGGAGCCGCTGCG
P K V D G L L H L D E L T R D A D L S A F V V F S S A A G V L G S A

CGGGCCAGGGCAGCTACCGCGCCCAACCGGGCGCTCGACGCGCTCGTGGCGCGGCTCACGCGCTCGGCCCTGCCGGCGTCTCCGTGGCGTGGGGCCT
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 16200
GCCCGTCCCGTCGATGCGCCGGGTTGCGCCGAGCTGCGCGAGCACCGCGGCGAGTGGCGAGCCGGACGGGGCCGAGAGGCACCGCACCCCGGA
G Q G S Y A A A N A A L D A L V R R R H A L G L P G V S V A W G L

Anexo I

GTGGGCCCGCCGAGCGGCATGACGGCGGGGCTCGACGAGGGCGACCGGGCGGGCTCGGCCGCTCCGGCGTCTCGAACTGCCACCGGACGACGCGCT
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 16300
CACCCGGCGGGCTCGCCGTACTGCCGCCCCGAGCTGCTCCCGCTGGCCCCGCGCCGAGCCGGCGAGGCCGAGGCTGACGGGTGGCTGTGCGCGA
W A A A S G M T A G L D E G D R A R L G R S G V L E L P T G R R A

CGCGCTGTTTCGACGCGCGACGCGCGACGACGCGGGTCTCTCGTGCCTGCTCTCGACACCGCCGCACTGCGCGGTCAGGCGGGCCGACGGCACCTGC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 16400
GCGCGACAAGCTGCGGGCTGCGCGCTGCTGCGCGCCAGGAGCACGGCCAGGAGGAGCTGTGGCGCGTGACGCGCCAGTCCGCCGGCTGCCGTGGGACG
R A V R R R D A R R R A V L V P V L L D T A A L R G Q A A D G T L P

CCCCGCTGCTCGGGGGCTGGTCCGCGCCCCCGCCCGCCGCGCCACGGCGGTGGCCCCGTGCGGGCCACGACGTTCCGCCACACCTCGGGGGCCGGCC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 16500
GGGGCGACGACGCCCCGACCAGGCGCGGGGGCGGGCGCGGTGCCGCCACCGGGGGCAGCGCCGGTGCTGCAAGCGGCTGTGGGAGCGCCCGCCGG
P L L R G L V R A P A R R A T A V A P V A A T T F A D T L A G R P

TGCCGAGGAACGCGAGCAGACGGTCTCGACCTGGTCTGCACGCAGGTGCGGGCGGTGCTCGGCTATGCGGCCGACACCGACGTCGACCCGGGCCGCGCG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 16600
ACGGCTCCTTGCCTGCTGTCAGGAGCTGGACCAGACGTGCGTCCAGCCCCGCCACGAGCCGATACGCCGGCTGTGGCTGCAGCTGGGCCGGCGCGC
A E E R E Q T V L D L V C T Q V G A V L G Y A A D T D V D P G R A

CTCGCGACCTCGGGTTCGACTCGTGCCTGACCTCGGTGGAACGCGCAACCGGCTCACCCGAGCCACCGGCTGCGGCTCGCCTCGACCTTCGTTTCGACC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 16700
GAGCCGCTGGAGCCCAAGCTGAGCGACTGGAGCCACCTTGACGCGTTGGCCGAGTGGCGTCCGTGGCCGACCCGAGCGGAGCTGGAAGCAGAAGCTGG
L G D L G F D S L T S V E L R N R L T A A T G L R L A S T F V F D H
ACP

ACCCGTCGGCCGCGGATCACCCGGTACCTGCTCACGCTGCTCCGAGCGACGAGCCGACCCGGGGACCGGCTCCTCGACCAGCTCGACCAACTGGA
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 16800
TGGGACGCGCGGCGCTAGTGGGCCATGGACGAGTGCACGAGGGCTCGTGTGCTCGGGCTGGGCCCTGGCCGACGAGCTGGTTCGAGCTGGTTGACCT
P S A A A I T R Y L L T L L P S D E P D P G T G V L D Q L D Q L E

GAGCGCGTGTGCTCGGCTCACCGACGACGGCACCGCCACCAAGGTGGACTGCGGCTGCGCGCCCTGCTCGCCAACGAGGACGACCCCGGGCCGATGAC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 16900
CTCGCGGACGAGCGGAGTGGCTGCTGCCGTGGCGGTGGTTCCACCTGGACGCCGACGCGGGACGAGCGGTTGACCTCGCTCGGGCCCGGCTACTG
S A L L A L T D D G T A T K V D L R L R A L L A N W S D A R A D D

GCGGAGGACGACGCGGAGCCGGCTTCGACGCGGCCACCGACGACGAACCTCTCGGCTACCTCGACAACGAACCTTGAGACCTGTAACCGGTGCGACCG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 17000
CGCCTCTGCTGCGCTCCGCCCCAAGCTGCGCCGGTGGCTGCTGCTTGAAGCCGATGGAGCTGTTGCTTGAACCTGAGGACACTGGCCAGCCTGGC
A E D D A E A G F D A A T D D E L F G Y L D N E L E T S *

AGGCGAACACGTCCGACCCCGTAAGGAGTACAGGCCAGATGGAGAACGAGGACAAGCTCCGCGATTACCTCAAGCGGGCCACCACCGATCTGCGCAAC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 17100
TCCGCTTGTCAGGCGTGGCGCATTCTCTATGTCGGTCTACCTCTGCTCTGTTGAGGCGCTAATGGAGTTCGCCCGGTGGTGGCTAGACGCGTTG

ACGCGCCGACGCTGCGGAGGTGAGGACCGGGAGCACGAGCCATCGCGATCATCGGGATGAGCTGCCGCTACCCCGGGCGGTGGCTCGCCGAGG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 17200
TGCGGGCTGCGGACGCGCTCCAGCTCTGGCCCTCGTGTGCGGTAGCGCTAGTAGCCCTACTCGACGGCGATGGGGCCCGCACCGGAGCGGGCTCC

ACCTGTGGCAGCTCGTCCGCCGCGGGGACGCGGTGGTGGAGTTCGCGCCGACCGCGGCTGGGACGCGGACGCTCTACGACCCCGACCCGGACCA
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 17300
TGGACACCGTCGAGCAGCGGCTGCCGCCCTGCGGCACCACCTCAAGGGCCGGCTGGCGCCGACCTGCGGCTGTGCGAGATGCTGGGCTGGGCTGGT
V A A R R R R G R R G G V P G R P R L G R R H A L R P R P G P
└─▶ *flvB* PstI

Anexo I

CGCGGGCACGTCCATCACCGGGGCACGGCGGTTTC**TCGAGGACGCGCGGGCTTCGACCCCGGCGTCTTCGGGGTCAGCCCGCGGAGGCCGCGGCGA**
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 17400
GCGCCCGTGCAGGTAGTGGCCCCGTGCCCAAAGGACGTCTGCGGGCCCCGAAGCTGGGGCCGCAAGAAGCCCGAGTCGGGCGCGCTCCGGCGCGCT
R G H V H H R G T A V S C R T P R A S T P A F F G V S P R E A A A M

TGGACCCGACGAGCGACTGCTCCTCGAGACGTCTGGGAGCGGTTTCGAGCGCGGGGCATCGACCCCGTGCAGGCCAAGGACACCAGGACCGGTGTCTT
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 17500
ACCTGGGCGTCTGCTGACGAGGAGCTCTGCAGCACCCCTCCGAAGCTCGCGCGCCCGTAGCTGGGGCACGTCCGGTTCCTGTGGTCTGGCCACAGAA
D P Q Q R L L L E T S W E A F E R A G I D P V Q A K D T R T G V F

CGCGGGCGTCATGTACAACGACTACGCGACGCGGCTGCCGACGGCGCCGAAGGGTTCGAGGGCTTCCTCGCCAACGGCAGCGGGGACGATCGCGTCC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 17600
GCGCCCGCAGTACATGTTGCTGATGCGCTGCGCCGACGGCGTCCGCGGGCTTCCCAAGCTCCCGAAGGAGCGGTTGCCGTCGCGCCCGTCGTAGCGCAGG
A G V M Y N D Y A T R L P Q A P E G F E G F L A N G S A G S I A S

GGCGGCATCGCCTACACCCTCGGCCCTCGTGGGGCCCGGTTGACGGTGGACACGGCGTGTCTCCTCGTCTGTCGCGCTGCACCTCGCGATGGTGGCGC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 17700
CCGGCGTAGCGGATGTGGGAGCCGGAGCAGCCCGGGGCCACTGCCACCTGTGCCGACGAGGAGCAGCGAGCAGCGGGACGTGGAGCGCTACCACCGCG
G R I A Y T L G L V **G P A V T V D T A C S S L** V A L H L A M V A L
KS

TGCGCCGTGAGGAGTGTCTCCCTCGCCCTGGCGGGCGCGCCACGTTTCATGGCCACGCGCGCACGTTTCGTTCGAGTACAGCAGGACGCGCCCTGTCCGT
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 17800
ACGCGGCACCTCCTCAGAGGGAGCGGACCGCCCGCGCGGTGCAAGTACCGGTGCGGCGCGTGAAGCAGCTCATGTCTGTCGCGCGGGACAGGCA
R R E E C S L A L A G G A T F M A T P R T F V E Y S R Q R A L S V

GGACGGCGCGTGAAGGCGTTCCTCCGACAGCGCCGACGGCACCGGTTGGGGCGAGGGCGTTCGCGATGCTCCTCGTGGAGCGCCTCCTCCGACGCCCGCGC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 17900
CCTGCCGGGACGTTCCGCAAGAGGCTGTGCGGGTGGCGGACCCCGCTCCCGCAGCCGTACGAGGAGCACCTCGCGGAGAGGCTGCGGGCGGGC
D G R C K A F S D S A D G T G W G E G V G M L L V E R L S D A R R

AACGGCCACCCGATCCTCGCGGTGGTCCGCGGCTCCGCGGTCACCCAGGACGGCGCGTCCACGGCCTGACCGCGCCCAACGGCCCTCCAGCAGACCG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 18000
TTGCCGGTGGGCTAGGAGCGCCACCAGGCGCCGAGGCGCCAGTTGGTCTGCGCGCAGGGTGCCGGACTGGCGCGGGTGGCGGGGAGGGTCTGCTGGC
N G H P I L A V V R G S A V N Q D G A S H G L T A P N G P S Q Q T V

TCATCAACCAGGCCCTGGCGAGCGCGGACTCACCCCGGCCAGGTCGACGCGCTCGAGGCGCACGGCACCGGCACCACCCTGGCGACCCCATCGAGGC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 18100
AGTAGTTGGTCCGGGACCGCTCGCGCGCTGAGTGGGGCCGGTCCAGCTGCGGCAGCTCCCGTGGCCGTGGTGGGAGCCGCTGGGGTAGCTCCG
I N Q A L A S A R L T P A Q V D A V E A H G T G T T L G D P I E A

GCACGCCCTCATCGCGACTACGGCGCCGAGCGCACCCCGGAGCGCCCGTGTGGCTGGGTTTCGCTGAAGTCGAACATCGGCCACGCCAGGGCGGGCC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 18200
CGTGGGGAGTAGCGTGGATGCCGCGGCTCGCGTGGGGCCTCGCGGGCGACACCACCAAGCGACTTCAGCTTGTAGCCGGTGGGGTCCGCCCGCGG
H A L I A T Y G A E R T P E R P L W L G S L K S N I G H A Q A A A

GGGGTGGCGGCATCATCAAGATGGTTCATGGCGATGCGGCACGGCACGCTGCCAGGACCCTGCACGTCGACGAGCCGAGCAGCCACGTCGACTGGTCCG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 18300
CCCCAGCCCGTAGTAGTTCTACAGTACCGCTACGCCGTGCCGTGCGACGGTCTGGGACGTGCAGCTGCTCGGCTGCTGGTGCAGCTGACCAGCC
G V G G I I K M V M A M R H G T L P R T L H V D E P S S H V D W S A

CCGGCACCGTGGCGCTGCTCACCGAGGCACAGCCGTGGCCGGAGACCGGCGAACCCGCGCCGCGCCGGTGTGTCTCTCGTTTCGGCGTACGCGGCACGAACAC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 18400
GGCCGTGGCACCCGCGACGAGTGGCTCCGTGTGCGGCACCGGCTCTGGCCGCTTGGCGCGGCGCGGCCACACAGGAGCAAGCCGAGTCGCCGTGCTTGTG
G T V A L L T E A Q P W P E T G E P R R A G V S S F G V S G T N T

CCACGTCATCTGGAGCAGGCACCCGCGCCCGGAGGAGACGCCCGCGCCGACCCGCGGCCACGACGACACCGAGCGTTCGGCGCCCGTCCAGC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 18500
GGTGCAGTAGGACCTCGTCCGTGGCGGGCGGGGCTCCTCTGCGGGCGGCGGTGGGCGCCGGGTGCTGCCTGTGGTTCGCGAGCCGCGGGGGCAGTGC
H V I L E Q A P A A P E E T P A A D P R P T T D T D A S A A P V T

CCCTGGGGCGTGTCCGCCCCGACCCCGACGCGCTGCGCGCACAGGCCGACCGACTCCGGTCCGCGTGGCGGCCCGCCGACTTCGACCCGGCGGGCA
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 18600
GGGACCCGCGACAGCGGGGCTGGGGCTGCGCGACGCGGTGTCGGCTGGCTGAGGCCAGCGGCACCCGCGCGCGGCTGAAGCTGGGCGCCCGT
P W A L S A R T P D A L R A Q A D R L R S A V A A A A D F D P A G T

CCGGCTTCTCCCTGGCCACGACCCGCTCCCTGTTCGAGCACCGCGCGTGTGGTGGCGGAGGACCGCGCCACCACCCTCGACCGGCTCACCGCTCTCGC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 18700
GGCCGAAGAGGGACCGGTGCTGGGCGAGGGACAAGCTCGTGGCGCGCACGACCACCGCTCCTGGCGCGGTGGTGGGAGCTGGCCGAGTGGCGAGAGCG
G F S L A T T R S L F E H R A V L V A E D R A T T L D R L T A L A

CGAAGCGCGGGCGCCGGGCGTCTGCTGGGCGAAGCCGCTCGCACCCGCTGGGCTTCTGTTCGCGGCCAGGGCTCGCAGCGCTCGGGATGGGT
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 18800
GCTTCCGGCGCCCGGGCCCGCAGCAGCACCCGCTTCGGCGCAGCGTGGCCGACCCGAAGGACAAGAGCCCGTCCCGAGCGTCGCGGAGCCCTACCCA
E G R G G P G V V V G E A A S H R L G F L F S G Q G S Q R L G M G

CGTGAAGTTCGCGGCTCGCTTCCCGGTGTTTCGCGGGCGCTGGACGAGGTCCTTGGCGAGTTCGATCCGTCGCTGCGTGAGGTGTTGTTTCGGTGAGGACG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 18900
GCACCTGAGCGCCGAGCGAAGGGCCACAAGCGCCGCGACCTGCTCCAGGAACGCTCAAGCTAGGCAGCCACGCACTCCACAACAAGCCACTCCTGC
R E L A A R F P V F A A A L D E V L A E F D P S V R E V L F G E D A

CGGACGCGTGAATGAGACCGGGGTGACGCAGCCCGTGTTCGCGGTTCGAGGTGGCGCTGTTCCGTTCTTTCGAGTCGTTGGGGCTGCGCCCCGACGT
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 19000
GCCTGCGGACTTACTCTGGCCCCACTGCGTGGGCGCAACAAGCGCCAGCTCCACCGCGACAAGGCAGAAGAACTCAGCACCCCGCACGCGGGGCTGCA
D A L N E T G V T Q P A L F A V E V A L F R L L E S W G V R P D V

GCTGGCGGGTCACTCGATCGGTGAGCTGGCCGCGGCATGTGGCCGGGTGTGGTTCGCTGGCGGATGCCGCCAAGTTCGTTGGCCGCGCGGGGCGCGCTG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 19100
CGACCCCGCAGTGAAGTACGACTCGACCGGCGCGGTACACCGGCCCCACACCAGCGACCCGCTACGGCGGTTCCAGCACCCGCGCGCCCCGCGGAC
L A **G H S** I G E L A A A H V A G V W S L A D A A K V V A A R G A L
ATa

ATGCAGGCGCTGCCGTCCGGTGGCGCGATGGTTCGCGATCCAGGCGTCCGAGGAAGAGTTCGCTCCGACGACCTGCCGGAGACCGTGGGCATCGCGGCCG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 19200
TACGTCCGCGACGGCAGGCCACCCGCTACAGCGCTAGGTCCGCGAGGCTCCTTCTCCAGCGCAGGCTGCTGGACGGCTCTGGCACCCGTAGCGCCGGC
M Q A L P S G G A M V A I Q A S E E E V A S D D L P E T V G I A A V

TGAACGGCCCGAACCTCCGTGGTTCGCTCCCGTGTTCGCGCTGATGTGAGGCGGTTCGCGGAGCGTTCGCGTGGAGGCGGGCCGCAAGGCACCCGGCTCAA
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 19300
ACTTGGCCGGCTTGAAGCACAGGAGGCCACAGCGGCGACTACAGCTCCGCCAGCGCTCGCAACCGCACTCCGCCGGCGTTCGCTGGGCGGAGTT
N G P N S V V V S G V A A D V E A V A E R W R E A G R K A T R L K

Anexo I

GGTCAGCCACGCGTTCCACTCGCCGCTCATGGACCCATGCTCGACGACTTCGCCGGGTCTGGAGGGCGTGTCTGACGAGACGCCCCAGATCCCCATC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 19400
CCAGTCGGTGCACAAGGTGAGCGGCGAGTACCTGGGGTACGAGCTGCTGAAGGCGGCCAGGACCTCCCGCACAGCATGCTCTGCGGGGTCTAGGGGTAG
V S H A F H S P L M D P M L D D F R R V L E G V S Y E T P Q I P I

GTGTGACGCTGACCGGTCTGCGGGCCACGGCCGAGGAGCTGACCTCCCCGGAGTACTGGGTGCGGCACGTGCGCGAGTCCGTGCGCTTCGCCGACGCGG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 19500
CACAGTGTGCGACTGGCCAGACGCCCCGGTCCCGGCTCCTCGACTGGAGGGGCTCATGACCCACGCCGTGCACGCGCTCAGGCACGCGAAGCGGCTGCGGC
V S T L T G L R A T A E E L T S P E Y W V R H V R E S V R F A D A V

TCGTGGCCCTGCGCGAGCAGGGCGCCGACGTGTTTCGTGGAGGTGCGCCCCGCGGAGTGTCTCCGGTCTCGGGCAGCTCACGGCCCCGGAGACCGCCTT
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 19600
AGCACCGGGACGCGCTCGTCCCGGCTGCACAAGCACCTCCAGCCGGGGCCCTCACGACAGGCCAGGCCGTCGAGTGCCTGGGGCTCTCGCGGAA
V A L R E Q G A D V F V E V G P G G V L S G L G Q L T A P E T A F

CGTACCGCCCTGCGCGGTGATCGCCCCGAGCTCGCGCTGACCATCGCCGTCGGCCACCTGCACGTGCGCGGTGTGCGGGTGAACCTGGCCCGCTGC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 19700
GCATGGCGGGACGCGCCACTAGCGGGCTCGAGCAGCGGACTGGTAGCGGAGCCGGTGGACGTGCACGCGCCACAGCGCCACTTGGACCGCGGACG
V P A L R G D R P E L V A L T I A V G H L H V R G V A V N L A A C

TTCGACGGCAGCGGCTCAGCGCTGGACCTGCCGACGTACCGTTCAGCGGAGCACTACTGGCTCGACGCGCCGACCCCGTCCGAGCGGGCGTTCGAGC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 19800
AAGCTGCCGTCGCGCGAGTCCGCGCACCTGGACGGCTGCATGCGCAAGGTGCGCTCGTGATGACCGAGTGCCTGGGGCAGCTCCGCCCGCAGCTCG
F D G S G A Q R V D L P T Y A F Q R E H Y W L D A D P V E A G V E R

GCATGTGCGCGGTGAGGCGCGGTTCGAGGAGCGGTGGAGGAAGGCGACCCCGCGACCTCGCCCGCACGCTGGGCGTGTCTGTCGGACGATCCGCTGAG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 19900
CGTACAGCGGCCAGTCCGGGCCAAGACCCTCCGGCACCTCCTCCGCTGGGGCGGCTGGAGCGGGCTGCGACCCGACAGCAGCCTGCTAGGCGACTC
M S P V E A R F W E A V E E G D P A D L A R T L G V S S D D P L S

CGCGGTGCTGCCTCGGCTGTCCGCTGGCGCGGGAAGCAGCGGGACCGGTCTGTGCGCGATGGCTGGCGGTACGGAGTGGAGTGGCGGCGGTGTCCGCC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 20000
GCGCCAGCAGCGGAGCCGACAGGCGCACCGCGCCTTCGTCGCCCTGGCCAGACAGCGGCTACCGACCGCCATGCTCACCTCACCGCCGGCCACAGGCGG
A V L P R L S A W R R K Q R D R S V A D G W R Y G V E W R P V S A

GGAGCCCGACAGTTGGACGCGCCTGGCTGGTCTGCTGCTGCTTCGACGGCGGAGGACCGCACCGAGTGGGTGCGTGCGGCGCTGGCCCGGACGGAGCTGACG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 20100
CCTCGGGCTGTCAACCTGCCGCGGACCGACAGCAGCGAAGCTGCCGCTCCTGGCGTGGCTCACCCACGACGCGCGGACCGGGCGTGCCTCGACTGC
G A R Q L D G A W L V V A S T G E D R T E W V R A A L A R D G A D V

TCCACGTCTGGACGTGGACCCGGAGACGGTTCGACTGGGCGGAGCAGTTGAGTGAACCTGCCCGGCTCGACGAGTCTGCTGCTGCTGGGGCTGGCGGA
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 20200
AGGTGACGAGCTGCACCTGGGCTCTGCCAGTACCCGCTCGTCAACTCACTTACGGGCGGAGTGCCTCAGCAGAGCGACGACCCCGACCGCCT
H V L D V D P E T V D W A E Q L S E L P A L D G V V S L L G L A D

CAGCGCGAGTTCGGCGGTGCCACGGGTGTGGCGGCCAGATCGGTCTGTTACCGCCCTGGGTGACGCGGACGTGTCCGCGCGCTGTGGTGCCTCAGC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 20300
GTCGCCGCTCAGCCGCCACGGGTGCCACACCGCCGGTGTAGCCAGACAATGCGCGGACCCACTGCGCCTGCACAGGCGCGGACACACGGAGTGC
S G E S A V P T G V A A T I G L L R A L G D A D V S A P L W C L T

TCGGGCGCGGTGGCGGCGCCGAGGGCGACGCGGTCCCCGGAATCGAGCAGTGCATGCTCTGGGGCTCGGGCGGGTGGCCCGCTGGAGCACCTGACC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 20400
AGCCCGCGCCACCGCCCGGCTCCCGCTGCGCCAGGGGCTTAGCTCGTACGCTACGAGACCCCGGAGCCGCGCCACCGCGGGACCTCGTGGGACTGG
S G A V A A A E G D A V P G I E Q S M L W G L G R V A A L E H P D R

Anexo I

GCTGGGGCGGACTGGTTCGACGTACCGGGCGCGCCGGACGAGACGACCGGTGACCTGCTGGCCACGGTCTCGCCGGTACGAGCGGGCAGGACCAACTCGC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 20500
CGACCCCGCTGACCAGTGCATGGCCGCCGCGCCTGCTCTGCTGGCCACTGGACGACCGGTGCCACGAGCGGCCATGCTCGCCGCTCTGGTTGAGCG
W G G L V D V P A A P D E T T G D L L A T V L A G T S G E D Q L A

BamHI
CCTGGGGGCGACGGATCTCGGCCGAGTTGGCGCCGCGCGCTGGGGGTGCCCGGAGAGGGCTGGTCCCGCGGGACGGTCTGGTGACC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 20600
GGACGCCCCGTGCCCTAGGAGCCGCGTCCAACCGCGGGCGGGACCCGCCACGGCGGCTCTCCCGACAGCGCGCGCCCTGCCAGGACCACTGG
L R G D G I L G R R L A P A P L G G A A G E G W S P R G T V L V T
KR

GGTGGCACCGGCGCCCTGGGCGGCCATGTGGCGCGTGGTTCGCGGCACACGGCGCGGAGCAGTCTGCTGGTGAGCCGTCCGGTGCCTCGGCCGAAG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 20700
CCACCGTGGCCGCGGACCCGCCGTACACCGCGGACCCAGCGCGTGTGCCGCGCTCGTGCAGCAGGACCACTCGGCAGCGCCACGGAGCCGCGCTTC
G G T G A L G G H V A R W V A A H G A E H V V L V S R R G A S A E G

SacI
CGCGCGCGCGCTGGAGGCCGAGCTACCGCGTCTCCACACGGCGGCATCGAGCGCTCCGCCCGCTCGCCGACCTCGACCCGGCGACCTCGCCGG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 20800
CGCGCCGGCGCGACCTCCGGTCTGAGTGGCGGCGAGGTTGTGCCCGCTAGCTCGCGAGGCGCGGGAGCGGCTGGAGCTGGGCCCGCTGGAGCGGCC
A A A L E A E L T A V V H T A G I E R S A A L A D L D P G D L A G

ATTCCGGCAGCTGCTGGCGGCAAGTTCGACGGCGCGGAACCTGCACGAACTCTGGGCGACACCCGCTGGACGCCTTCGTGCTGTTCTCGTCCATC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 20900
TAAGCCGCTGCACGACCGCGGTTCAGCTGCCGCGCCTTGGACGTGCTTGGAGACCCGCTGTGGGGCGACCTGCGGAAGCAGACAAGAGCAGGTAG
F G D V L A A K V D G A R N L H E L L G D T P L D A F V L F S S I

GCCGGAGTCTGGGGCAGCGCGGCCAGGCCCGCTACGGCGCGCCAACGCCTACCTCGACGCCCTGGCCGAGCACCGCAGGGCGCGGGCCTGCCCGCA
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 21000
CGGCCTCAGACCCCGTCCGCCCGGTCCGCCGATGCCGCGCGTTCGGATGGAGCTGCGGGACCGGCTCGTGGCTCCCGCGCCCGGACGGCGCGGT
A G V W G S G G Q A A Y G A A N A Y L D A L A E H R R A A G L P A T

CCGCGGTCCGCTGGGGCCGTTGGGCCGACGGCGGCATGGTCCGCGACCGCGTGGCGACGTGGCCCGCAACTGCACCGCAGGGCCCTGACCACGATGGC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 21100
GGCGCCAGCGCACCCCGGCACCCGGTTCGCCCGTACCAGCGCCTGCGGCCACGCCCTGCACCGCGCGTTCAGTGGCGTCCCGGACTGGTGTACCG
A V A W G P W A D G G M V A D A G A D V A A Q L H R R G L T T M A

CCCCGACGCGCGCTCGCCGGCTCGGCGCGGCCATCGACCACGGCGACGGACAGATCACCATCGCCGACGTCGACTGGTCCCGGTTCCGGGGCACCTTC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 21200
GGGGTTCGCGCGAGCGGCCGAGCCGCGCGCGGTAGCTGGTGCCTGCTAGTGGTAGCGGTGCAGCTGACCAGGGCCAAGCGCCCGTGGAG
P D A A L A G L G A A I D H G D G Q I T I A D V D W S R F A G T F

ATGGCGCAGCGCCCAGCCCGTCTCGGTGAACCTCCCGACATCCGGGCCGCGTTCGACACCGCTCCGGCCGAGCAGGCCACCTCGGCCCTCGCCAGC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 21300
TACCGCTCGCGGGTTCGGCGAGGACCACTTGGGGGCTGTAGGCCCGGCAAGCTGTGGCGAGCGCGTTCGTCGGTGGAGCCGGGAGCGGGTTCG
M A Q R P S P L L G E L P D I R A A F D T A P A E Q A T S A L A Q R

GCCTGGCCGGTCTGGACGAGCCGAGCAGGAGCGGCAGCTCACCGAGCTGGTCCGGTACGAGGCGCGCGGTGCTCGGCCATGCGTCCGGCGACGCCTT
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 21400
CGGACCGGCCAGACCTGCTCCGGTCTGCTCCCGTCCGAGTGGCTCGACCAGGCCATGCTCCGCCGCGCCACGAGCCGGTACGACGCGCGTTCGGGAA
L A G L D E A E Q E R Q L T E L V R Y E A A A V L G H A S G D A F

CTCGGCCACCCGCGCCTTCGCGACCTCGGCTTCGACTCCTTGACGGCCGTCGAACTGCGCGCCCGCCTCAGCGAGGCGACCGCGGTGGCCCTGCCCGCC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 21500
GAGCCGGTGGGCGGGAAGCGCTGGAGCCGAAGCTGAGGAAGTCCCGGACGCTTACGCGCGGGCGGAGTTCGCTCCGCTGGCCGACCGGGACGGCGG
S A T R A F R D L G F D S L T A V E L R A R L S E A T G V A L P A
ACP

Anexo I

GCCGTGGTCTTCGACTACCCGACGGCCGCGGCTGGCCACCTATCTGCGGACCGAACTGTCCGGAGCTGCCGCCACCCCACTGCCGCGGGCCG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 21600
CGGCACCAGAAGCTGATGGGCTGCCGGCCGCGGACCCGGTGGATAGACGCCTGGCTTGACAGGCCTCGACGGCGGGTGGGGGTGACGGCGCCGCGGC
A V V F D Y P T A A A L A T Y L R T E L S G A A A A T P T A A A A V

TGGCGGGTGCAGCCGACGAGCCGATCGGATCGTGGCGATGTCGTGCCGGTCCCCGGCGGGTGTCCAGCCCGGAGGACCTGTGGCAGGTGCTGCGGTC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 21700
ACCGCCACGCTGGCTGCTCGGCTAGCGCTAGCACCCTACAGCACGGCCAAGGGCCGCGCACAGGTCGGGCTCCTGACACCCGTCCACGACGCCAG
A G A T D E P I A I V A M S C R F P G G V S S P E D L W Q V L R S

GGGCGTGGACGCGGTGACCGACTTCCCGACCGACCGGGCTGGGACGTCGAGGCGCTGTACGACCCGAGGGCGGGCGTCCGCACGTCTACTCCCGC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 21800
CCCGCACCTGCCCACTGGCTGAAGGGCTGGCTGGCCCCGACCTGCAGCTCCGCGACATGCTGGGCTCCCGCGCCCGCAGCCGTGCAGCATGAGGGCG
G V D A V T D F P T D R G W D V E A L Y D P E G A G V G T S Y S R

ATGGGTGCCCTTCCTCGATCGCGTGGCGGACTTCGACCCCGCTCTTCGGGATCTCGCCGCGGAGGCCCTCGCCATGGACCCGACGAGCGGTGCTCC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 21900
TACCCACGGAAGGAGCTAGCGCACCCCTGAAGCTGGGGCGGAGAAGCCCTAGAGCGGCGGCTCCGGGAGCGGTACCTGGGCGTCTCGCCGACGAGG
M G A F L D R V A D F D P A L F G I S P R E A L A M D P Q Q R L L L

TGGAGACGTCGTGGGAGGCGTTCGAGCGGGCGGCATCGACCCGACGTCGCTGCGCGGCGAGCCGGTCCGGTGTGTTTCGTCCGACCAACGGCCAGGACTA
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 22000
ACCTCTGCAGCACCTCCGCAAGCTCGCCCGCCCTAGCTGGGCGTCAGCGACGCGCCGTCGGCCAGCCACACAAGCAGCCGTGGTTGCCGTCCTGAT
E T S W E A F E R A G I D P Q S L R G S R S G V F V G T N G Q D Y

CACCCGGTGTGATGAGCCGCTCAGGACACCGTTCGACCGGCTACTTCGGCACCCGCGAGCGGGCGAGCGTCATCTCCGGCCGCTCTCGTACACGTTCCGGC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 22100
GTGGGCGACTACTACCGGCGAGTCTGTGGCAGCTGCCGATGAAGCCGTGGCCGTCGCGCCGCTCGCAGTAGAGCCGGCGGAGAGCATGTGCAAGCCG
T R L M M A A Q D T V D G Y F G T G S A A S V I S G R L S Y T F G

CTCGAAGGCCCGCCGTCACCGTGGACACGGCGTCTCGTTCGCTCGCTCGTCCCTGCACCTGGCCGTGCAGGCCCTGCAGAAGGGCGAGTCCGAGCTGG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 22300
GAGCTTCCGGCCGCGAGTGGCACCTGTGCCGACGAGCAGCAGCAGCAGCGGGACGTGGACCCGCGACGTCGGGACGTCCTCCCGCTCACGCTCGACC
L E **G P A V T V D T A C S S S** L V A L H L A V Q A L Q K G E C E L A
KS

CGCTCGCGGGCGGTGTGACGGTGTGTCGACGCCCGCGGCTTCGTGGAGTTCAGCCGCCAGCGGGGCTCTCGGTGGACGGCCGGTGAAGCGTTCGG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 22400
GCGAGCGCCCGCCACTGCCACTACAGCTGCGGGCGGCAAGCACTCAAGTCGGCGGTGCGCGCCGAGAGCCACTGCCGGCCACGTTCCGCAAGCC
L A G G V T V M S T P A A F V E F S R Q R G L S V D G R C K A F G

CGCGGACCCGACGGCACCGGCTGGGGCGAAGGCGTCGGCATGCTCCTCGTCGAGCGCCTTCCGACCCGAGCGCAACGGCCACCAGGTGCTCGCCGTC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 22500
GCGCTGCGGGTGCCTGGCCGACCCCGCTTCCGACGCCGTCGAGGAGCAGCTCGCGGAAAGGCTCGGGTTCGGTTCGCGGTGGTCCACGAGCGGCAG
A D A D G T G W G E G V G M L L V E R L S D A Q R N G H Q V L A V

GTGCGGGCTCGGGGTCAACCAGGACGGCGCTCCAACGGCTCACCGCGCCAACGGCCCTCACAGCAGCGGGTTCATCCGCGCCGCGTGGCGACCG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 22600
CACGCGCCGAGCCGAGTTGGTCTGCCGCGAGTTGCCGAGTGGCGGGTTCGGGGAGTGTGTCGCCCAGTAGGCGGGCGGACCGCTGGC
V R G S A V N Q D G A S N G L T A P N G P S Q Q R V I R A A L A T A

Anexo I

CGGGCCTGGCGCCGTCGATGTGGACGCGGTGCGAGGCACACGGCACGGGTACGAAGCTGGGCGACCCGATCGAGGCCAGGCGCTCCTCGCCACCTACGG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 22700
GCCCGGACCGCGGCAGGCTACACCTGCGCCAGCTCCGTGTGCCGTGCCCATGCTTCGACCCGCTGGGCTAGCTCCGGTCCGCGAGGAGCGGTGGATGCC
G L A P S D V D A V E A H G T G T K L G D P I E A Q A L L A T Y G

CCAGGAACGTGCCGGTGGCCGCTGTGGCTCGGCTCGATCAAGTCGAACATCGGCCACACCCAGGCTGCCGCCGGTGTGCGCCGGTGTGATCAAGATGGTC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 22800
GGTCCTTGACGCGCCACTCGGCGACACCGAGCCGAGCTAGTTCAGCTTGTAGCCGGTGTGGGTCCGACGGCGGCCACAGCGGCCACACTAGTTCTACCAG
Q E R A G E P L W L G S I K S N I G H T Q A A A G V A G V I K M V

ATGGCGATGCGCCACGGCGTACTGCCGCGGACGCTGCACGCGGACGAGCCGTCGCCGCACGTGGACTGGTCCGCCGGTGGGTCGAGGTGCTGTGCGGAGG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 22900
TACCGCTACGCGGTGCCGCATGACGGCGCCTGCGACGTGCGCCTGCTCGGCAGCGGCTGCACCTGACCAGGCGGCCACGCCAGCTCCACGACAGCCTCC
M A M R H G V L P R T L H A D E P S P H V D W S A G A V E V L S E A

CCCGCGGTGGCCTGCGCGGACCGGCCCTGGCGTGGCGTGTGTCGTCGTTCCGGTTCAGCGGCACCAACGCGCACGCCATCATCGAGCAGGCCCCCGT
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 23000
GGGCGCGCACCGGACGGCGCCTGGCCGGGACCGCACGGCCACACAGCAGCAAGCCCAAGTCCCGTGGTTCGCGGTGCGGTAGTAGCTCGTCCGGGGGCA
R P W P A A D R P W R A G V S S F G F S G T N A H A I I E Q A P V

GGTCGTACGGGAGGACGTCCCGGAGCCGCTGCCGGTGGCGCGGTGTCGCCGTCCCGTGGCTCCTTTCGGCCAAGACCCCGAGGCGCTGCGGGAGCAGGCCCGG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 23100
CCAGCATGCCCTCCTGCAGGGCCTCGGCGACGGCCAGCCGCGCCAGCAGGGCACCGAGGAAAGCCGGTTCGCGGGTCCGCGACGCCCTCGTCCGGGCC
V V R E D V P E P L P V G A V V P W L L S A K T P E A L R E Q A R

CGTCTCACGTCCATGTGGACGCGGCCCTGAGCTGTCCGTGCCGGACGTGGGCCGTGCGCTGGTGTCTCGCGGGCGGTGCTCGAACACCGAGCCGTGG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 23200
GCAGAGTGCAGGATACACCTGCGCCGGGACTCGACAGGCACGGCCTGCACCCGGCACGCGACCACAAGAGCGCCCGCCACGAGCTTGTGGCTCGGCACC
R L T S Y V D A A P E L S V P D V G R A L V F S R A V L E H R A V V

TCGTGCGCCCGGACGCGCCCGGCTTCCAGTGGGCCCTGGCCGCGCTCGCCGAAGGGCGTGGCGGTTCGGGTGTGCGGGTGGGTGAGGCGGCACCGCACCG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 23300
AGCAGCGCGGCTGCGCGCGCCGAAGGTTCAGCCCGGACCGCGCGAGCGGCTTCCCGCACCGCCAAGCCACAGCGCCACCCACTCCGCCGTGGCGTGGC
V A A D A A G F Q S G L A A L A E G R G G S G V A V G E A A P H R

GTTGGGCTTCCGTGTTCTCGGGACAGGGGTGCGAGCGCCTGGGGATGGGCCGTGAACCTGGAGGCCCGTTCGCCGGTGTTCGCCGAGGCGCTGCACGAGGTG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 23400
CAACCCGAAGGACAAGAGCCCTGTCCCAGGCTCGCGGACCCCTACCCGGCACTTGACCTCCGGGCAAGGGCCACAAGCGGTCCCGGACGTGCTCCAC
L G F L F S G Q G S Q R L G M G R E L E A R F P V F A E A L H E V

CTTGGGAGTTCGACCCGTCGGTGGCTGCGGTCCTGTTCCGGCGGACCGCCGACGCGCTGAACGAGACCGGCGTGACGAGCCCGGCTGTTCCGGGTCG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 23500
GAACGCTCAAGTGGGACGACCGCACGCCAGGACAAGCCGCGCCTGCGGCTGCGGACTTGTCTGGCCGACTGCGTCCGGCGCAACAAGCGCCAGC
L A E F D P S V R A V L F G A D A D A L N E T G V T Q P A L F A V E

AGGTGGCGCTGTTCCGCTCCTTGAATCCTTGGGTGTCCGGGCCGACGTGCTCGCCGGCCACTCGATCGGTGAGCTGGCCGCCGCGCATGTGGCCGGGGT
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 23600
TCCACCGCACAAGGCGGAGGAACTCAGGAACCCACAGGCCCGGCTGCACGAGCGCGGCTGAGCTAGCCACTCGACCGGCGCGCTACACCGGCCCA
V A L F R L L E S L G V R A D V L A **G H S** I G E L A A A H V A G V
ATa

GTGGTTCGTCGGGACCGCCGAAGTTCGTCGGCGCGGGGTGCGCTGATGACGGCGCTGCCGTCGGTGGCGGATGGTGGCGATCCAGGCCACGGAG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 23700
CACCAGCGACCGCTGCGCGGTTCCAGCACCGCGCGCCACCGGACTACGTCGCGACGGCAGGCCACCGCGCTACCACCGCTAGGTCCGGTGCCTC
W S L A D A A K V V A A R G A L M Q A L P S G G A M V A I Q A T E

Anexo I

GAAGAGGTTCGCGTCCGAGGGCTGCCGGAGACCGTGGGCATCGCTGCCGTCAACGGGCCCTCGTCCGTGGTTCGTTTCGGGTGTTGCCGCCGATGTCGAGG
 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 23800
 CTTCTCCAGCGCAGGCTCCCGGACGGCCTCTGGCACCCGTAGCGACGGCAGTTGCCGGGAGCAGGCACCAGCAAAGCCACAACGGCGGCTACAGCTCC
 E E V A S E G L P E T V G I A A V N G P S S V V V S G V A A D V E A

CGGTGCCGAGCGCTGGCGTGGCGGGCCGGAAGGTGTCGCCCTCAAGGTGACCCAGCGTTCCACTCGCCGCTGATGGACCCGATGCTCGACGAGTT
 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 23900
 GCCAGCGGCTCGGACCCGACGCCGCCCGCCGCTTCCACAGGGCGGAGTTCCAGTCCGTGCGCAAGGTGAGCGCGACTACCTGGGCTACGAGCTGCTCAA
 V A E R W R A A G R K V S R L K V S H A F H S P L M D P M L D E F

CCGTGCGGTGGTTCGAGGGCATCGAGTGGCGGAACCCGAGATCCCGATCGTGTGCGACGCTGACCCGACAGCTCGCCACGGCCGAGGAGTTGGCCTCCCGG
 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 24000
 GGCACGCCACCGACTCCCGTAGCTCACGCCGCTTGGGCTCTAGGGTAGCACAGCTGCGACTGGCCTGTGAGCGGTGCCGGTCTCTAACCCGAGGGGC
 R A V V E G I E C G E P E I P I V S T L T G Q L A T A E E L A S P

GAGTACTGGGTGCGGCACGTCGCGAGTCCGTGCGCTTCCACGACGCGCTCGTGGCGTGC GCGAGCAGGGCACCGATGTGTTCGTCGAGGTGCGCCCGG
 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 24100
 CTCATGACCCACGCCGTGCACGCGCTCAGGCACGCGAAGGTGCTGCGGCAGCACCCGACGCGCTCGTCCCGTGGCTACACAAGCAGCTCCAGCCGGGCC
 E Y W V R H V R E S V R F H D A V V A L R E Q G T D V F V E V G P G

GCGGCACGCTGTCCGCTCTCGGCCTGGGCAGCGCGCCGACGCGGCGTTCCCTGCCGCGTGC GCGCGGACAGGCCCGAGGAGAACGCGCTCCTCGGCAC
 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 24200
 CGCCGTGCGACAGGGCAGAGCCGGACCCGCTGCGCGGGCTGCGCCGCAAGGACGGGCGGACGCGCGCTGTCCGGGCTCCTCTTGC GCGAGGAGCCGTG
 G T L S A L G L G S A P D A A F L P A L R A D R P E E N A L L G T

GGTTCGCCGCGTGCACACGCAGGGCGTGGCGTGCAGTGGGCGCGTACTTCCGCGGCGGGCGCGCCCGCCCGTTCGACCTGCGGACCTATGCCTTCCAG
 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 24300
 CCAGCGGCGGACGTGTGCGTCCCGCACGGCCAGCTGACCCGGCGCATGAAGCGCCCGCCCGCGGGCGGGGCGAGCTGGACGGCTGGATACGGAAGGTC
 V A A L H T Q G V P V D W A A Y F A G G G A R P V D L P T Y A F Q

CGCGAGCGCTACTGGTGGACGCCGCTCCGTGCGATGACGGCGCCGAGCGCATGTGCGCGGTGCGAGGCGGGTCTGCGGAGGTGCTCGAGGAAGCGGACA
 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 24400
 GCGCTCGCATGACCGACCTGCGGCGCAGGCAGTACTGCGCGGCTGCGGTACAGCGGCCAGCTCCGCGCCAAGACCTCCAGCAGCTCCTTCCGCTGT
 R E R Y W L D A A S V D D G A E R M S P V E A R F W E V V E E G D I

TCGCCGATCTCGCCCGACGCTCGAGGTGTCGTCGGACGATCCGCTGAGCGCGGTGCTGCCTCGGCTGTCCGCTGCGCGTGGCGCCGGAAGCAGCGGGACCGGTC
 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 24500
 AGCGGCTAGAGCGGGCGTGCAGCTCCACAGCAGCTGCTAGGCGACTCGCGCCACGACGGAGCCGACAGGCGCACCCGCGCCTTCTGCGCCTGGCCAG
 A D L A R T L E V S S D D P L S A V L P R L S A W R R K Q R D R S

TGTCGCCGATGGCTGGCGGTACGGAGTGGAGTGGCGGCCGGTGTCCGCGGAGCCGACAGCTGGACGGCGCCTGGTGGTGTGCTTCGACGGGCGAG
 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 24600
 ACAGCGGCTACCGACCGCCATGCCTCACCTACCGCCGGCCACAGGCGGCTCCGGGCTGTGCGACTGCGCGGACCGACCGAGCAGGAAGCTGCCCGGCTC
 V A D G W R Y G V E W R P V S A G A R Q L D G A W L V V A S T G E

GACCGCACGAGTGGGTGCGTGGCGGCTGGCCCGACGGAGCTGACGTCGCGTCTGGACGTGGACCCGAGACGGTTCGACTGGGCGGAGCAGTTGA
 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 24700
 CTGGGCTGGCTACCCACGCACGCCGCGACCGGGCGCTGCCTCGACTGCAGGCGCAGGACCTGCACCTGGGCCTTGCAGCTGACCCGCTCGTCAACT
 D R T E W V R A A L A R D G A D V R V L D V D P E T V D W A E Q L S

GTGAACTGCCCGCGTTCGACGGGTCGTCGCTGCTGGGGCTCGCCGACGTGGGCGACTCGGCTGTGCCGACGGGTGTCGCGGGCACGGTTCGATCTGCT
 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 24800
 CACTTGACGGGCGGAGCTGCCCCAGCAGAGCGACACCCCGAGCGGCTGCACCCGCTGAGCCGACACGGCTGCCACAGCGCCCGTCCAGCTAGACGA
 E L P A L D G V V S L L G L A D V G D S A V P T G V A G T V D L L

Anexo I

GCGGGCGCTCGGTGACGCGGGTGTGGAGGCGCCGTTGTGGTGTCTCACGTGCGGGCGGGTGGCGGGCGCCAGGGCGACGTCGTCGCCGGGTTTCGAGCAG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 24900
CGCCCCGAGCCACTGCGCCACACCTCCGCGGCAACACCACAGAGTGCAGCCCGCCACC GCCCGGCTCCCGTGCAGCAGCGGCCAAGCTCGTC
R A L G D A G V E A P L W C L T S G A V A A A E G D V V A G F E Q

TCGATGCTGTGGGGTCTGGGCGGGTGGCCGCCCTGGAGCAGGCGGTCGTTGGGGGGTCTGGTCGACGTACCCGCGGCGCCGGACGAGACACCGGCG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 25000
AGCTACGACACCCAGACCCGGCCACC GGCGGGACCTCGTCCGGCCAGCAACCCCGCAGACCAGTGCATGGGCGCCGCGCCTGCTCTGCTGGCCGC
S M L W G L G R V A A L E Q A G R W G G L V D V P A A P D E T T G D

ATCAGCTGGCCGCGTCCTTGCCGGGACGAGTGGCGAGGACCAGGTCGCGCTGCGCGGCGACGGATCCCGCCGACGGCTGGCGCCCGCGCCGCTCAG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 25100
TAGTCGACCGGCGGAGCAACGGCCCTGCTACCGTCTGTTCCAGCGCGACGCGCCGCTGCCCTAGGAGCCGGCGTCCGACCGGGCGCGCGGAGTC
Q L A A V L A G T S G E D Q V A L R **G D G I L G R R L A** P A P L S
KR

CGCGACCGGCTCCGGTCCGGCTGGACGCCGCGCGCACGGTCCTGGTGACCGGTGGCACCGGCGCCCTGGGCGGTTCATGTGGCGCGCTGGCTGGCGGCG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 25200
GCGCTGGCCGAGGCGGAGGCCGACCTGCGGCGCGCCGTGCCAGGACCCTGCCACCCTGGCCGCGGGACCCGCCAGTACCCGCGGACCGACCGCCGC
A T G S G S G W T P R G T V L V T G G T G A L G G H V A R W L A A

AGCGGTGCGGACGCTCGTACTGGCGGGGCGTCGCGGCCCGCGCCGAGGGCGCCGAGGAACTGGAGGCGGAGCTCGTAGCGGGGCGCGCGCTCA
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 25300
TCGCCACGGCTCGTGCAGCATGACCGCCCGCAGCGCCGGCGCGGCTCCCGCGCTCCTTGACCTCCGCCTCAGCAGCTCGCCCGCGCGCGCAGT
S G A E H V V L A G R R G P A A E G A E E L E A E L V E R G A R V T

CCATCGCCGCTGCGACGTGACGGACAAGGAATCGGTCCGGGCGCTCCTGGAGAGCGTCCCCGCGCTCACCGCCGTCGTCACACGGCGGGCATCGAGCG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 25400
GGTAGCGGCGGACGCTGCACTGCCTGTTCTTAGCCAGGCCCGCAGGACCTCTCGCAGGGGCGGAGTGGCGGCAGCAGGTGTGCCCGCCGTAGCTCGC
I A A C D V T D K E S V R A L L E S V P A L T A V V H T A G I E R

CTCGGCCACCCTCGCCGAGCTGGACCGGGACACCTGGCCGGTGTGCTCGCGGCGAAGGTCGACGGCGCGCGGAACCTGCACGAGCTCTGGGCGACACG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 25500
GAGCCGGTGGGAGCGGCTCGACCTGGCCCTGCTGGACCGGCCACACGAGCGCCGCTTCCAGCTGCCGCGCCCTTGGACGTGCTCGAGGACCCGCTGTGC
S A T L A E L D R D D L A G V L A A K V D G A R N L H E L L G D T

CCGCTGGACGCCTTCGTGCTTCTCGTCCATCGCGGGAGTCTGGGGCAGCAGCGGGCAGGGCGGTCACGCGCCGCCAACGCCTACCTGGATTCCCTGG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 25600
GGCGACCTGCGGAAGCACGACAAGAGCAGGTAGCGCCCTCAGACCCCGTCTGTCGCCGTCCCGCGCATGCGCCGGCGGTTGCGGATGGACCTAAGGGACC
P L D A F V L F S S I A G V W G S S G Q G A Y A A A N A Y L D S L A

CCGAGCACCGCAGGGCCGCGGCCTGCCCGCACCGCGGTCGCGTGGGGGCGGTGGGGCGGCGGCGGCGCATGGTCGCCACCGCCGAGCAGACCGCGCTGCT
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 25700
GGCTCGTGGCGTCCCGGCGCCCGGACGGGCGGTGGCGCCAGCGCACCCCGGACCCCGCCCGCCGTACCAGCGGTGGCGGCTCGTCTGGCGCGACGA
E H R R A A G L P A T A V A W G P W G G G G M V A T A E Q T A L L

GCAGCGCTGGGGCCTGACGACGCTGGTTCCGGCGGTGGCGATCGCCGGTCTGGCCACGGCCTTGGAGCGCGGTGACGGGAACGTCGTCGTCGCGGATGTC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 25800
CGTCGCGACCCCGACTGCTGCGACCAAGGCCGCCACCGCTAGCGGCCAGACCGGTGCCGGAACCTCGCGCCACTGCCTTGCAGCAGCAGCGGCTACAG
Q R W G L T T L V P A V A I A G L A T A L E R G D G N V V V A D V

Anexo I

GACTGGTCCCGGTTTCGCGGGCACCTTCATGGCGCAGCGCCCCAGCCCGCTCCTCGGTGAACTGCCCGAGGTGCAGGCCGCGTTTCGGGACCTCTGCGGCCG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 25900
CTGACCAGGGCCAAGCGCCCGTGGAAGTACC CGCTCGCGGGGTTCGGGCGAGGAGCCACTTGACGGGCTCCACGTCCGCGCAAGCCCTGGAGACCGCCGG
D W S R F A G T F M A Q R P S P L L G E L P E V Q A A F G T S A A E

AGCAGGGCACGTCGGCGCTCGCGCAGCGGCTGGCCGGTCTGGACGAGGCGGAGCAGGAGCGGCAGCTCACCGAGCTGGTCCGGTACGAGGGCGGCCGCTGT
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 26000
TCGTCCCGTGCAGCCGCGAGCGCTCGCCGACCGGGCAGACCTGCTCCGCCTCGTCTCGCCGTCGAGTGGCTCGACCAGGCCATGCTCCGCCGGCGACA
Q G T S A L A Q R L A G L D E A E Q E R Q L T E L V R Y E A A A V

CCTGGGCCACGTCCTCCGGCGACGCGTCTCCTCCGCGACCCGCGCCTTCCGTGACCTCGGCTTCGACTCCCTGACGGCCGTCGAACTGCGCGCCCGCCTCAGC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 26100
GGACCCGGTGCAGAGGCCGCTGCGCAAGAGGCGCTGGGCGCGGAAGGCACTGGAGCCGAAGCTGAGGGACTGCCGGCAGCTTGACGCGGGCGGAGTGC
L G H V S G D A F S A T R A F R D **L G F D S L T A** V E L R A R L S
ACP

GACGCCACCGCCTCGCCCTGCCGGCCACGCTGGTCTTCGACTACCCGACGGCCGCGCTCCTGGCCGCTACCTGCGCAGCGAACTGTGCGGTGCCAGTG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 26200
CTGCGGTGGCCGAGCGGGACGGCCGGTGCACAGAACTGATGGGCTGCCGGCGCAGGACCGGGGATGGACGCGTGCCTTGACACGCCACGGTAC
D A T G L A L P A T L V F D Y P T A A V L A A Y L R S E L C G A S G

GGGCACCCACGGCCTCGGCTCGGCTCGCGTCTCCGGTGTGAGCGACGAGCCGATCGCGATCGTGTGATGTCGTGCCGGTACXXXXXXXXXXXXXXXXXX
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 26300
CCCGTGGGTGCCGGAGCCGAGGCCAGCGCCAGAGGCCACACTCGCTGCTCGGCTAGCGCTAGCACAGCTACAGCACGGCCATGXXXXXXXXXXXXXXXXXX
A P T A S A P V A V S G V S D E P I A I V S M S C R Y

XX
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 26400
XX

XX
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 26500
XX

XXXXXXXXXXGATCGGCTTCGAGACCCGCGCTCACCGTCGACACGGCGTGCTCCTCGTCGCTGGTCGCGCTGCACCTGGCGGTGCAGGCGCTGCGGTCC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 26600
XXXXXXXXXXCTAGCCGAAGCTCTGGGGCGGAGTGGCAGCTGTGCCGCACGAGGAGCAGCGACCAGCGCAGCTGGACCGCCACGTCGCGCAGCCAGG
V V H D R L R D **P A V D T A C S S** L V A L H L A V Q A L R S
KS

GGAGAGTGGGAGCTGGCGCTCGCGGGCGGCTGACGGTGTGTCACGCCCGGGCGGTTTCATCGAGTTCAGCCGCCAGCGGGCCTGTCCCTCGACGGCC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 26700
CCTCTCACGCTCGACCAGGAGCGCCCGCCGACTGCCACTACAGTGGCGGCGCGCAAGTAGCTCAAGTCGGCGGTGCGCCGGACAGGGAGCTGCCGG
G E C E L A L A G G V T V M S T P G A F I E F S R Q R G L S L D G R

GGTGAAGGCCTTCGCGCCCGCCGCGACGGCACCAGCTGGGGCGAGGGCGTCGGCATGCTCTTCGTCGAGCGCCTCTCGGACGCCAGCGCAACGGGCA
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 26800
CCACGTTCCGAAGCGGGCGGGCGGCTGCCGTGGTCGACCCCGCTCCCGCAGCCGTACGAGAAGCAGCTCGCGGAGAGCCTGCGGGTTCGCGTTCGCCGT
C K A F A A A A D G T S W G E G V G M L F V E R L S D A Q R N G H

CCCCGTCTCGGGTATCCGCGGCTCCGCGGTGAACCAGGATGGCGCCTCCAACGGCCTCACCGCACCGAACGGCCCTCACAAACAGCGGTATCCGC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 26900
GGGCGAGGAGCGCCACTAGGCGCGGAGGCGCCACTTGGTCTACCGGGAGGTTGCCGGAGTGGCGTGGCTTGCCGGGAGTGTGTGTCGCGCAGTAGGCG
P V L A V I R G S A V N Q D G A S N G L T A P N G P S Q Q R V I R

Anexo I

CAGGCCCTGGCCAGCGGGCCTGGCCGCGCCGACGTCGACGCCGTCGAGGCGCACGGCACCGGCACGAAGCTCGGCGACCCGATCGAGGCCAGGCGC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 27000
GTCCGGACCGGTTCGCGCCCGGACCGGCGCCGGCTGCAGCTGCGGCAGCTCCGCGTCCGCTGGCCGTGCTTCGAGCCGCTGGGCTAGCTCCGGGTCCGCG
Q A L A S A G L A A A D V D A V E A H G T G T K L G D P I E A Q A L

TCCTCGCCACCTACGGCCAGGACCGCCCGCCGACCAGCCCTGTGGCTCGGGTCGGTCAAGTCCAACATCGGCCACACGAGGCCCGCCGGTGTGGC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 27100
AGGAGCGGTGGATGCCGGTCTTGGCGGGGCGGCTGTCGGGACACCGAGCCAGCCAGTTTCAGGTTGTAGCCGGTGTGCGTCCGGCGCGGCCACACCG
L A T Y G Q D R P A D Q P L W L G S V K S N I G H T Q A A A G V A

GGGCGTGATCAAGATGGTCATGGCGATGCGGGCGGGCGTGCTGCGCGCACCCCTGCACGTGGACGAGCCGTCGCCGCACGTCGACTGGTCCGCGGGCGCG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 27200
CCCGCACTAGTTCTACCAGTACCGCTACGCCCGCCGACGACGGCGCGTGGGACGTGCACCTGCTCGGCAGCGGGCGTGCAGCTGACCAGGCGCCCGCGC
G V I K M V M A M R A G V L P R T L H V D E P S P H V D W S A G A

GTGGAAGTGTACCGATCAGCGTGCCTGGCCCGCCGTTGGACCGCCGTTGGCTGCCGGTGTGTCTGCTCGGTTTCAGCGGCACCAACGGCGACGCCA
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 27300
CACCTTGACGAGTGGTAGTTCGACGACCGGGCGGACCTGGCCGCGACCGCACGCCACACAGCAGCAAGCCCAAGTCGCCGTGGTTCGCGGTGCCGT
V E L L T D Q R A W P A V D R P W R A G V S S F G F S G T N A H A I

TCATCGAGCAGGCCCGCGGCGACGCCCGGAGGCGGACGGGCGGACGGAGCCCGCGTCTGCCGTGGGTCTCTCGGCCCGGTCCGGCGAGGGCGCT
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 27400
AGTAGCTCGTCCGGGGCGCGCTGCGGGCGGCTCCGCCGTCGCCGCTGCCGCGCAGGACGGCACCCAGGAGAGCCGGGCCAGGCCGCTCCGCGA
I E Q A P A A T P A E A A G P T E P G V L P W V L S A R S G E A L

GCGGCAACAGGCCCGCGGCTCCGGGACSacI**GAGCTCGACCACAAGCGGACGTGCCGGTGCCTCGCCACGCTCGCCACAGCTGCGGTCCCGGCCGGCGTTCGAC**
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 27500
CGCGTGTTCGCGCGGGCCGAGGCCCTGCTCGAGCTGGTGTTCGCGCTGCACGGCCAGCGGTGCAGCGGGTGTGCGACGCCAGGGCCGGCCCAAGCTG
R Q Q A A R L R D E L D H K A D V P V A D V A H T L R S R P A F D

CACCGCGCGTGTGGTGGCCCTCGACCGGAGGACTTCTCAGCGGCTCACCGCCCTGGCCGACGGCCGAGGCCCGAACATCGTCACCGGGCGCG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 27600
GTGGCGGGCACGACCACCGGAGCTGGCGCTCCTGAAGGAGTCGCCGAGTGGCGGGACCGGCTGCCGGCGCTCCGGGGCTGTAGCAGTGGCCGCGCC
H R A V L V G L D R E D F L S G L T A L A D G R E A P N I V T G A A

CCGGCAGCGGGCACCGGCACGGTGTTCGTGTTCCCGGCCACGGCACGCGAGTGGTTCGGCATGGCCCTCGAACTCCTGGAGTCTCGCCGGTGTTCGG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 27700
GGCCGTTGGCGCGGTGGCCGTGCCACAAGCACAAGGGGCGGTTGCCGTGCGTCACCCAGCCGTACCGGGAGCTTGAGGACCTCAGGAGCGGCCACAAGGC
G D R G T G T V F V F P G H G T Q W V G M A L E L L E S S P V F R

CGCGCGGATGCAGGAGTGCAGCGAGGCGCTCGGCGAGTACGTGACTGGTTCGCTCCTCGACGCGTCCAGGACCCCGCCCTGATGGAGCGCTGGACGTC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 27800
GCGCGCTACGTCTCACGCGCTCCGCGAGCCGCTCATGCAGCTGACCAGCGAGGAGTGCGGCAGGTCTGGGGCGGGACTACCTCGCGGACCTGCAG
A R M Q E C A Q A L G E Y V D W S L L D A V Q D P A L M E R L D V

GTCCAGCCCGCCCTCTTCGCGGTGATGGTCTCGCTGGCCGCCACCTGGCAGTTCGTGGGGCGTTCGAGCCGAGCGCCGCTCTGGGGCACTCGCAGGGCGAGA
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 27900
CAGGTGGGGGGGAGAAGCGGCACTACCAGAGCGACCGGCGGTGGACCGTGCAGACCCCGCAGCTCGGCTCGCGGCAGGACCCCGTGGAGCTCCCGCTCT
V Q P A L F A V M V S L A A T W Q S W G V E P S A V L **G H S** Q G E I
ATp

TCGCGGGCGCTGCGTTCGCGGGCGCCCTGCCCTGAAGGACGCGGCGCCCTCGTCTGCTGCGCAGCAAGGCGCTGACCGCGCTCGCGGGACAGGGCGG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 28000
AGCGCCCGGACGACGCGCCCGGAGCGGGACTTCTGCGCCGCGGAGCAGCAGGACGCGTTCGCGACTGGCGGAGCGCCCTGTCCCGCC
A A A C V A G A L P L K D A A R L V V L R S K A L T A L A G Q G G

Anexo I

CATGGTGTCCGTCGCGCTCGACGTCGACGCCGTACGGGACCTCATCGGCCCTGGGACGGACAGCTCGACATCGGCGTGGTGAACGGCCCCGCTCGGTC 28100
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
GTACCACAGGCAGCGCGAGCTGCAGCTGCGGCATGCCCTGGAGTAGCGCGGGACCCTGCCTGTCGAGCTGTAGCCGCACCACTTGCCGGGGCGGAGCCAG
M V S V A L D V D A V R D L I A P W D G Q L D I G V V N G P A S V

GTCGTGTCCGGCGACGCCCGCGCTGGACGAACTGGTGGACCTCTGCGTCGAGCAGGGCGTGTGGGCGGGCGGGTCCAGGCGGACTACGCCCCGCACT 28200
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
CAGCACAGGCCGCTGCGGGCGCGCAGCTTGCTTGACCACCTGGAGACGCAGCTCGTCCCGCACACCCGCGCCGCCAGGTCCGCCTGATGCGGGGCGTGA
V V S G D A A A L D E L V D L C V E Q G V W A R R V Q A D Y A P H S

CCCACCACGTCGAGCGGATCGAGGCGGAACTCCTGGCGGGCGTCGCGGAGGTGACGCCGCGCTCCGGAGACGTGCCGTACTTCTCGCCGCTGACCGGCGG 28300
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
GGGTGGTGCAGCTCGCTAGCTCCGCTTGAGGACCGCCCGCAGCGCTCCACTGCGGCGCGAGGCCCTTGCACGGCATGAAGAGCGGCGACTGGCCGCC
H H V E R I E A E L L A G V A E V T P R S G D V P Y F S P L T G G

CTGGTGGACACCCGCGAGCTGGACGCCGCTACTGGTACCAGCCTGCGCCAGACCCTGTCAGTTCCGCCACTCCGTCTCCGCACTGGTGGCGGGCGGC 28400
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
GACCGACTGTGGGCGCTCGACCTGCGGCGCATGACCATGGCGTCGGACGCGGTCTGGCACGTCAAGCGGCTGAGGCAGAGCGGTGACCACCGCCGCCCG
W L D T R E L D A A Y W Y R S L R Q T V Q F A D S V S A L V A A G

TTCAGCACCTTCGTCGAGGTGAGCCCGACCCGGTCTCGTGGCGGGTCTGGCGAACTGGTTCGGCGCCGACGGCGTGGTTCGTCGGCTCGCTGCGGCGTG 28500
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
AAGTCGTGGAAGCAGCTCCAGTCGGGCGTGGGCCAGGAGCACCGGCCAGAACCCTTGACCAGCCGCGGCTGCCGCACCAGCAGCCGAGCGACGCCGCAC
F S T F V E V S P H P V L V A G L G E L V G A D G V V V G S L R R G

GGGACGGCGGCCCGCCGGCTCATGACCTCGGCGGCCACGCTGTTACGCGCGGGCGCCGGTGGAGTGGCCGCTCCCGGCAGCGCAGCGCGTCGACCT 28600
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
CCCTGCCGCCGGGCGGGCCGAGTACTGGAGCCCGGTCGACAAAGTGCAGCGCCCGGCCACCTCACCGCGAGGGCCGTCGCGTCGCGCAGCTGGA
D G G P A R L M T S A A T L F T R G G R V E W P L P A A Q R V D L

GCCGACGTACCCCTTCCAGCAGGAGCGCTTCTGGCCACGGCGGCACACCCGCTCCGGCGACGTGACCTCGGCCGGCTCGGCCGGTGGACCACCC 28700
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
CGGCTGCATGGGAAGGTCGTGCTCGGAAGACCGGTCGCCCGTGTGTGGCGCAGGCCGCTGCACTGGAGCCGGCCGAGCCGGCCACCTGGTGGGG
P T Y P F Q H E R F W P T A A H T A S G D V T S A G L G R V D H P

GTCCCTCGTCTGTCACCGAACTCGCCGACTCCGAAGGCCCTCGTGTTCAGCGGCCGACTGTCCCTGCGCACCACCCGTGGCTCACCGACTACCGCGTAC 28800
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
CAGGAGCAGCGACAGTGGCTTGAGCGGCTGAGGCTTCCGGAGCACAAAGTGCAGCGGCTGACAGGGACGCGTGGGTGGGCACCGAGTGGCTGATGGCGCATG
V L V A V T E L A D S E G L V F S G R L S L R T H P W L T D Y R V L

TCGGCACCGCGTGTCCCGGCGCGGCTACGTGGACCTCGCGCTGCGCGCGGTCGACACGTCGGCTGCGTGTCCCTGGACGACCTCGTCTCGGAGGC 28900
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
AGCCGTGGCGGCACGAGGGGCGCGCCGATGCACCTGGAGCGCAGCGCGGCCACTGGTGCAGCCGACGCACAGGGACCTGCTGGAGCAGGACCTCCG
G T A V L P G A A Y V D L A L R A G D H V G C V S L D D L V L E A

GCCCTGGTGTGCCCGAGCGGGACGGCGTGCAGGTGCGGCTCAGCCTCGACGGGCGGACGCGGGCGGGCGGGTTCACCGTGGACTCCCGAGGG 29000
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
CGGGGACCACGACGGGCTCGCCCTGCGCACGTCCACGCCGAGTCGGAGCTGCCCGGCTGCGCCGCCGGCCGCCCAAGTGGCACCTGAGGGCTCCC
P L V L P E R D G V Q V R L S L D G P D A A G R R A F T V D S R G

GACCGGGCGGCGCTGGACCGCCACGCCACCGGCACCCTGGCGGCGGACGCGGAAACCCGTCACCGACGACCTCGTCCGCTGGCCCGGACCGAGGCCG 29100
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
CTGCGCCCGCGGACCTGGGCGGTGCGGTGGCCGTTGGGACCGCCGGCTGCGGCTGGGCGAGTGGCTGCTGGAGCAGCGGACCGGCGGCTGGCTCCGGC
D A G G A W T R H A T G T L A A D A E P V T D D L V A W P P T E A E

Anexo I

AACCGGTCGGGATCGACGAGCACTACGCCACCATCGCCGCTCCGGCCTCGACCACGGCCCCGCCTTCCAGGGCCTGCGCGCCGCTGGCGCAAGGGCGA
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 29200
TTGGCCAGCCCTAGCTGCTCGTGATGCGGTGGTAGCGGGCGGAGCCGGAGCTGGTGCCGGGGCGGAAGGTCCCGGACGCGCGGCGGACCGCGTTCCCGCT
P V G I D E H Y A T I A A S G L D H G P A F Q G L R A A W R K G D

CGAGGTGTTCCGCCAAGTGGTGTGCTGCCGGACGGCCTCACGGACGAGGGCTTCGGCCTGCACCCGGCCCTCCTGGACGCCCGCTGCACACGGCCGGGCTC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 29300
GCTCCACAAGCGGCTTACCACGACGGCCTGCCGGAGTGCCTGCTCCCGAAGCCGGACGTGGGCCGGGAGGACCTGCGGCGCGACGTGTGCCGGCCCGAG
E V F A E V V L P D G L T D E G F G L H P A L L D A A L H T A G L

GGGGT^{SacI}**GAGCTCGGACGGGCTCCCGTTCTCGTGGGCCGGGTGCGGCTGCTCGCCACCGGGCGGACGGTGTGCGGGTGCAGCCCGTGGCGGTGG**
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 29400
CCCCACTCGAGCCTGCCCCAGGGCAAGAGCACCCGGCCCCACGCCGACGAGCGGTGGCCGCGCTGCCACGACGCCACGCCGACTGCGGGCAGCCGCACC
G V S S D G L P F S W A G V R L L A T G A T V L R V R L T P V G V D

ACGCCGTGGCGGTGCTCGTCGCCGACGGCTCCGGCGCACCGGTGCGCCGCTCGACCAGTTGAGCCTGCGGCCCTCTCGGCCGACCAGTCAACCTCGC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 29500
TGCGGCACCGCCACGAGCAGCGGCTGCCGAGCCGCGTGGCCAGCGCGGCAGCTGGTCAACTCGGACGCCGGGAGAGCCGGCTGGTTCGAGTTGGAGCG
A V A V L V A D G S G A P V A A V D Q L S L R P L S A D Q L N L A

CCGGTCCGCGCACCGGACTCGTGTTCGAGGTGAAGTGGGTCCCGTACGGGACGCGCTCCCGACACCGGCACCTGGGCGGTGCTCGGGCGGACACC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 29600
GGCCAGGCGCGTGGCGCTGAGCGACAAGCTCCACTTCAACCCAGGGGCATGCCCTGCGGCGAGGGCTGTGGCCGTGGACCCGCCACGAGCCCGCTGTGG
R S A H R D S L F E V K W V P V R D A A P D T G T W A V L G G D T

GCCCGCTCGCCACGCGCTCACCGGGGGGGGAGCCGGTCGACGAGCACGGGACCTGGCCGCGCTCCGCGCCGCGTTCGACGGGGGTGCGCCCGTGC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 29700
CGGGCGGAGCGGGTGGCGAGTGGCGCCGCCGCTCGGCCAGTGTGCTGCGGCTGGACCGCGCGAGGCGGGCGGCGAGCTGCCCCACGCGGGCAGC
A R L A H A L T A A G E P V D E H G D L A A L R A A V D G G A P V P

CCCGCGCGTGGCTGTGACGCCCGGCGAGGGCGCGACGCCCGCGCGCGGACCCCGGACACCTCGACCTGCTCCAGCAGTGGCTCGCCGACGACAG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 29800
GGGCCGCGCACCGACTGCGGGCCGCTCCCGCGCTGCTGGGGCGCGCCGCTGGGCGCTGTGGGAGCTGGACGAGGTGTCACCGAGCGGCTGCTGTC
R R V A V T P G E G A T T P R A A T R D T L D L L Q Q W L A D D R

GTTCGCGGACGCGCAGCTCGTGTCTCACGAGCGGCGGGTGGCGGCGAGGCCCGACGAGGACCTCCCGACCTGGTCCGCGCCCCCGGTGGGGCCTG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 29900
CAAGCGGCTGCGGCTCGAGCACGAGGAGTGTCTCGCCGCGCCACCGCGTCCCGGGCTGCTCCTGCAGGGCCTGGACCAGGCGGGGGCGCACCCCGGAC
F A D A Q L V L L T S G A V A A G P D E D V P D L V R A P A W G L

GTGCGGTGCGGCGAGTCGGAACACCCGGAGCGGTTCTGCTCCTCGTACGCTCGACGCCACGAGAAGTCTGCTGCGCGGCTGCCCGCGCGGTGGCCACCG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 30000
CACGCCAGCCCGTCAAGCTTGTGGGCTCGCCAAGCAGGAGCAGCTGCAGCTGCGGGTGTCTTTCAGCGACGCGCCGGACGGGGCGGCCACCGGTGGC
V R S A Q S E H P E R F V L V D V D A H E K S L R G L P A A V A T G

GGGAGCCGACGCTCGGATCAGGAACGGCGCCGCTCGGTGGCCAGGCTGGCGGCTCGCGGCTCCGGCCGACGCGCCGAGGCCGACCACGACCCGA
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 30100
CCCTCGGCTCGAGCCTAGTCTTCCCGCGGCGGACCCCGTCCGACCGCGCCGAGCGCCGAGGCGGCTGCGCGGGCTCCGGCTGGTGTGCTGGGCT
E P Q L A I R N G A A S V A R L A R L A A P A D A P E A D H A P E

Anexo I

GCCCGGGCGTTACCGCCGACAGCACCGTCTGATCACCGGCGGTCTCGGCCTACTCGGCGGCCTACCGCACGCCACCTCGTCACCGCCACGGCGTG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 30200
CGGGCCCCGCAAGTGGCGGTGTCTGTGGCAGGACTAGTGGCCGCCAGAGCCGCATGAGCCGCCGAGTGGCGTGGAGCAGTGGGCGGTGCCGCAC
P G A F T A D S T V L I T G **G L G V L G G L T A** R H L V T R H G V
KR

GGCAAGCTGGTCTCACCGCCCGCGCGGCCCGGCCACGCCGGGTGCGGAGGAACGTGCCGGGAACGGAGGAGTCGGGGGCACAGGTCACGATCGCCG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 30300
CCGTTTCGACCAGGAGTGGCCGGCGGCCCGGCCGTTGCGGCCACGCCTCCTTGACACGGCCCTTGACCTCCTCAGCCCCGTGTCCAGTGTAGCGGC
G K L V L T G R R G P A T P G A E E L C R E L E E S G A Q V T I A A

CGTGCGACGTGAGCGACCGCGACCGGTGGCGGCGCTCTTGGCCGAGCACCCCGTGACCGCCGTGCGTGCACACCGCCGGTGTCTGGACGACAGCGTGGT
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 30400
GCACGCTGCACTCGTGGCGTGGCCACCGCCGCGAGAACC GGCTCGTGGGGCACTGGCGGCAGCACGTGTGGCGGCCACAGGACCTGTGTGCACCA
C D V S D R D A V A A L L A E H P V T A V V H T A G V L D D S V V

CGAGTCCCTGACCCCGGAGCACGTGACGCGGTGTTCGGCCCCAAGGGCGACGCCGCGCTGCACCTGCACGAGCTGACCCCGACCTGGACCTGTGCGGC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 30500
GCTCAGGGACTGGGGCCTCGTGCAGCTGCGCCACAAGCCGGGTTCCCGCTGCGGCGGCAGCTGGACGTGCTCGACTGGGCGCTGGACCTGGACAGCCGC
E S L T P E H V D A V F G P K G D A A L H L H E L T R D L D L S A

TTCTGCTGTCTCCCGCTCGCCGGTGTCTTCGGCAGCCCGGCCAGGGCAACTACGCGCGGGCAACAGCTTCTGGACGCGCTCGCCAGCACCGCC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 30600
AAGCAGGACAAGAGGCGGCAGCGGCCACAGAAGCCGTGGGCGCCGTTCCCGTGTGATGCGCCCGCTGTGCGAAGGACCTGCGCGAGCGGGTGTGGCGG
F V L F S A V A G V F G S P G Q G N Y A A A N S F L D A L A Q H R R

GGGCGCGGACTGCCCGGACGTGCTGCGCTGGGCGCTGTGGGCGGAGGCCAGCGGCATCACCGGGCACCTCGACGACGCTGACGTACAGCGGATGGC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 30700
CCC GCGGCCGTGACGGGCGTGCAGCGAGCGGACCCCGGACACCCGCTCCGGTTCGCGTAGTGGCCCGTGGAGCTGTGCGACTGCATGTGCGCTACCG
A A G L P A T S L A W G L W A E A S G I T G H L D D A D V Q R M A

GCGCGCGGCATGGCCCCGTGTCTCCCGGACGGCATGGCACTCTTCGACGCGGCGCGCGGACCGGGCGCGGACGGCGGTGCCCGCCGATCGAC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 30800
CGCGCCCGCTACCGGGGCGACAGGAGGCGGCTGCCGTACCGTGTGAGAAGCTGCGCCGCGCGCTGGCCGCGGCGTGGCCACGGGCGGCGCTAGCTG
R G G M A P L S S A D G M A L F D A A R A T G A A T A V P A A I D

CTGGCGGCGTGGGGCCACCCGAGCAGGTACGGCCGTGCTGCGGCGCTGGTGCCACGCCCGCCGAGGGCCGCGGCTCCGGCGACGGTCCCG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 30900
GACCGCCGCGACGCCCGGTTGGGCTCGTCCATGCCGGCGACGACGCGCCGACCAGGGTGCGGGCGGGCGTCCCGCGCGCCAGGCCGCTGCCAGGGC
L A A L R A H P E Q V R P L L R G L V P T P A R R A A R S G D G P G

GCACGGTCTCGCTCGCCAGCGGCTCGCGGGCCTCACCGAGGGAGCGGGACCGCACCTCGTCCACCTCGTCCGACGCCACCGCGGGCGTGTGCGG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 31000
CGTGCCAGAGCGAGCGGGTCCCGAGCGCCGAGTGGCTCCGCTCGCCCTGGCGTGGGAGCAGGTGGAGCAGGCGTCCGTTGGCGCCGGCACGAGCC
T V S L A Q R L A G L T E A E R D R T L V H L V R S H T A A V L G

CTTCGCGGGCCGAGGCCGTGAGGCGGAACGGGCTTCAAGGAACCTCGGATTCGACTCGCTCACCGGCGTGGAACTGCGCAACAGGCTCGACGCCGAG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 31100
GAAGCGGCCCGCCTCCGGCAGTCCGCCTTGCCCGAAGTTCCTTGAACCTAAGCTGAGCGAGTGGCCGACCTTGACGCGTGTCCGAGCTGCGGCTC
F A G P E A V E A E R A F K E **L G F D S L T G V E** L R N R L D A E
ACP

GTCGGCACGGACTGCCCGCCACCCCTGGTCTTCGACCACCCGACCCCGGGCGCACTCGCGGCCACCTGCGGACGACGGTGTCTGCCGACGACCGCACGC
 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 31200
 CAGCCGTGCGCTGACGGGGCGTGGGACCAGAAGCTGGTGGGCTGGGGCCCGCTGAGCGCCGGGTGGACGCCCTGTGCCACGAGCGGTGTGGCGTGCG
 V G T R L P A T L V F D H P T P A A L A A H L R T T V L A D D R T P

CCGCCCGGAGGTGTCTGGCGAACTGAACAAGCTGACGGTCAACATCTCCAAGCTCGACCCGAACGACGCGTGGTTCGGAGACATCCGGCTGGGGCTCCG
 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 31300
 GGCGCGGGCTCCACGAGCCGTTGACTTGTTCGACTGCCAGTGGTAGAGTTTCGAGCTGGGCTTGTCTGCGCGACCAGCCTCTGTAGGCCGACGCCGAGGC
 A A E V L G E L N K L T V T I S K L D P N D A L V G D I R L R L R

GTCACTCTGTCCACCTGGGACGACAAGGAGCCGGAGAGCGGGCGGGCGACCTCGACTCGGCAACGCTCGACGACGTGTTCAACATCATCGACGAGGAA
 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 31400
 CAGTGAGGACAGGTGGACCCTGTGTTCCTCGCCCTCTCGCGCCCGCGTGGAGCTGAGCCGTTGCGAGCTGCTGCACAAGTTGTAGTAGCTGCTCCTT
 S L L S T W D D K E P E S A A G D L D S A T L D D V F N I I D E E

CTTGAGTCTGAACCGCAGGTCCGGACTCTACCAAAGGGTAACAAAGCACATGTGAACGACGACAAGTCTCGACTACCTGAAGCGCGTCACCGC
 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 31500
 GAACTCAGGACTTGGCGTCCAGGCCCTGAGAGTGGTTTTCCCATTTGTTTCGTGTACAGCTTGTCTGCTTCCAGGAGCTGATGGACTTCGCGCAGTGGCG
 L E S *

RBS

GGACCTGCGGCAGACCAAGCAACGGTGCAGGAGT**CAGGAGAAGG**CCGCCGAGCCGATCGCGATCGTGGC**ATGAGCTGCCGGTTC**CCCGGGCGGC**TCT**
 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 31600
 CCTGGACGCCGCTCTGGTTCGTTGCCGACGCGCTCCAGCTCCTTCCGGCGGCTCGGCTAGCGCTAGCAGCCGTA**CTCGACGCCAAGGGGCCCGCAG**
 M S C R F P G G V
▶ **flvC**

GAGTCGCCCGAGGACCTGTGGCGTCTGGTTCGAGTCCGGTACGGACGCGGTGTTCGGACTTCCCAGCCGACCCGCGGCTGGGACGTGGAGTCGCTCTACGACC
 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 31700
 CTCAGCGGGCTCCTGGACACCGCAGACCAGCTCAGGCCATGCCTGCGCCACAGCCTGAAGGGCTGGTGGCGCCGACCCTGCACCTCAGCGAGATGCTGG
 E S P E D L W R L V E S G T D A V S D F P T D R G W D V E S L Y D P

CGGATCCCAGCCGCTCCGGCACCCACTACAGCCGCAAGGGCGCCTTCTCGACCCGCGCGGAGTTCGACTCCGGCTTCTTCGGGATCTCGCGCGCGCA
 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 31800
 GCCTAGGGCTGGCGAGGCCGTTGGTGGATGTCGGCGTTCCCGCGGAAGGACCTGGCGCGCGCCTCAAGCTGAGGCCGAAGAAGCCCTAGAGCGGCGCGCT
 D P D R S G T T Y S R K G A F L D R A A E F D S G F F G I S P R E

GGCCTTAGCGATGGACCCGACGAGCGGCTGTCTCGAGACGTCGTGGGAGGCGTTCGAGCGGGCGGCATCGACCCGCACACCCTGCGCGGACCGGG
 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 31900
 CCGGAATCGCTACCTGGGCGTCGTCGCCGACGACGAGCTCTGCAGCACCCCTCCGCAAGCTCGCCCGCCGTAGCTGGGCGTGTGGGACGCGCCGTCGGCC
 A L A M D P Q Q R L L L E T S W E A F E R A G I D P H T L R G S R

TCCGGCGTGTTCGTCGGCACCAACGGCCAGGACTACGGGCCCTGTTGATGGCCGCGATGGAGACCGTCGAGGGCTACGGGGCACCGGCATCTCGGCGA
 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 32000
 AGGCCGCACAAGCAGCCGTGGTGGCGGTCCTGATGCGCCGGGACAACTACCGGGCTACCTCTGGCAGCTCCCGATGCGCCCGTGGCCGTAGAGCCGCT
 S G V F V G T N G Q D Y A A L L M A A M E T V E G Y A G T G I S A S

CGGTCACTCCGGCCGCTCTCGTACAGTTCGGCCTCGAAGGCCCGGCGTACCCGTGGACACGGCTGCTCGTCTGCTGGTTCGCCCTGCACCTCGC
 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 32100
 CGCAGTAGAGCCCGGAGAGCATGTGCAAGCCGGAGCTTCCGGGCCGCGAGTGGCACCTGTGCCGACGAGCAGCAGCACCAGCGGGACGTGGAGCG
 V I S G R L S Y T F G L E G P A V T V D T A C S S S L V A L H L A
 KS

GGCGCAGGCCCTGCACAAGGGCGAGTGCACATGGCCCTCGTCGGCGGCGTCACGGTGATGTCGACACCCGGCTCCTTCGTGGAGTTCAGCCGCCAGCGC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 32200
CCGCGTCCGGGACGTGTTCCCGCTCACGCTGTACCGGGAGCAGCCGCCGAGTGCCTACTACAGCTGTGGGCCGAGGAAGCACCTCAAGTCGGCGGTTCGCG
A Q A L H K G E C D M A L V G G V T V M S T P G S F V E F S R Q R

GGCCTCTCGGTGGACGGCCGCTGCAAGGCGTTCGCCGAGGCCGCCGACGGCACCCGGCTGGGGCGAGGGCGGGCATGCTCCTCGTGGAGCGCCTCTCGG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 32300
CCGGAGAGCCACCTGCCGGGACGTTCCGCAAGCGGCTCCGGCGGCTGCCGTGGCCGACCCCGCTCCCGCGCCGTACGAGGAGCACCTCGCGGAGAGCC
G L S V D G R C K A F A E A A D G T G W G E G A G M L L V E R L S D

ACGCCCAGCGCAACGGCCACAAGATTCTCGCGGTTCGTCGGCGGCTCGCGGGTCAACCAGGACGGCGCTCCAACGGCCTCACCGCCCCGAACGGCCCGTC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 32400
TGCGGGTTCGCGTTGCCGGTGTCTAAGAGCGCCAGCACGCGCCGAGCCGCCAGTTGGTCTGCCGCGCAGGTTGCCGGAGTGGCGGGGCTTGCCGGGCGAG
A Q R N G H K I L A V V R G S A V N Q D G A S N G L T A P N G P S

GCAGCAGCGCGTCATCCGCCAGGCCCTGGAGAGCGGGCCCTCGCCCCGACCGACGTCGACGCCATCAGAGGCACACGGCACCCGGCACCCCTCGGCGAC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 32500
CGTCGTCGCGCAGTAGGCGGTCCGGGACCTCTCGCGCCCGGAGCGGGGCTGGCTGCAGCTGCGGTAGCTCCGTGTGCCGTGGCCGTGGTGGGAGCCGCTG
Q Q R V I R Q A L E S A G L A P T D V D A I E A H G T G T T L G D

CCGATCGAGGCCCAGGCCCTCCTCGCCACCTACGGCCAGGGCCGACGGCGGAGCCCTGTGGCTCGGCTCCGTGAAGTCCAACGTTGGCCACACCCAGG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 32600
GGCTAGCTCCGGGTCCGGGAGGAGCGGTGGATGCCGGTCCCGGCGTCCCGCTCGGGGACACCCGAGCCGAGGCACTTCAGGTTGCAACCGGTGTGGGTCC
P I E A Q A L L A T Y G Q G R S G E P L W L G S V K S N V G H T Q A

CCGCCCGCGGTGTTCGGGGCGTGATCAAGATGGTTCATGGCGATGCGCCACGGCGTGTGCCGAGACGCTGCACGTTGGACCAGCCCTCGACGCACGTCGA
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 32700
GGCGGGGCCACAGCGCCCGACTAGTTCTACCAGTACCGCTACCGGTGCCGACGACGGCGTCTGCGACGTGCACCTGGTGGGAGCTGCGTGCAGCT
A A G V A G V I K M V M A M R H G V L P Q T L H V D Q P S T H V D

CTGGTTCGGCGGGCGCGTCCGAACCTCCTCACGGAGGCCCGGGCGTGGCCGAGACCGGCCACCTGCGCGCGCGCGGTGTGTCTCCTTCGGCGTGAGCGGC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 32800
GACCAGCCCGCCGCGCAGCTTGGAGAGTGCCTCCGGGCCGGCACCCGGCTCTGGCCGGTGGACGCGGCGCGGCCACACAGCAGGAAGCCGCACTCGCCG
W S A G A V E L L T E A R P W P E T G H L R R A G V S S F G V S G

ACCAACGCGCACGCCATCATCGAGCAGGCCCCCGCGCTCGAGCAGGAGGAGGCCCGGAGCCGGTACGCGACGGCCTGGTGCCTGGGTCCTCTCGGCC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 32900
TGGTTGCGCGTGCCTAGTAGCTCGTCCGGGGGCGGAGCTCGTCCCTCCGGGGCCTCGGCCATGCGCTGCCGACCACGGGACCCAGGAGAGCCGGG
T N A H A I I E Q A P A V E Q E E A P E P V R D G L V P W V L S A R

GGAGCCCCGAGGCGCTGCGCGTCCAGGCCGAGCGGCTCCGCTCCACGTCGCGGACCTCGACCCCCAGGACGTCGGCTTCTCGCTCGCCACCACCCGGGC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 33000
CCTCGGGGCTCCGCGACGCGCAGGTCGGCTCGCCGAGGCGAGGTTGACGCGGCTGGAGCTGGGGTCTGCAGCCGAAGAGCGAGCGGTGGTGGGCCCG
S P E A L R V Q A E R L R S H V A D L D P Q D V G F S L A T T R A

GGCGCTGGAACACCGTGCCTCGTCAFCGCGGACGACCACGGGTCCCGCTGACCGCCCTCACGGCACTGGCCGAGGCCAGGCAGACCCCGGGCCTGGTG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 33100
CCGCGACCTTGTGGCACGGCAGCAGTAGCGCTGCTGGTGGCCAGGCGGACTGGCGGGAGTCCCGTGACCGGCTCCGGTCCGTCTGGGGCCCGGACCAC
A L E H R A V V I A D D H G S R L T A L T A L A E A R Q T P G L V

ACGGGCTCCACCGGCAGGGCCTGCTGGGCTTCTGTTCCTCGGGCCAGGGCTCGCAGCGCCTCGGCATGGGGCGTGAACCTGTCCGAGCGCTTCCCGGTGT
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 33200
TGCCCGAGGTGGCCGTCGCCGGACGACCCGAAGGACAAGAGCCCGGTCCCGAGCGTCGCGGAGCCGTACCCCGCACTTGACAGGCTCGCGAAGGGCCACA
T G S T G S G L L G F L F S G Q G S Q R L G M G R E L S E R F P V F

TCGCCGAGGGCTGGACGAGTCTTGGGAGTTCGACCCGTCGGTACGTGAGGTGCTGTTCGGTGGGGACCGGGACGCCCTGAACGAGACGGGAGTGAC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 33300
AGCGGCTCCGCGACCTGCTCCAGGAACGCCTCAAGCTGGGCAGCCATGCACCTCCACGACAAGCCACCCTGCGCCTGCGGGACTTGCTCTGCCCTCACTG
A E A L D E V L A E F D P S V R E V L F G G D A D A L N E T G V T

GCAGCCCGCGCTGTTCGGGTTCGAGGTTCGCGCTGTTCGCCCTCCTTGAGTCTGGGGTGTGCGCCCCGACATGCTGGCCGGGCACTCCATCGGGAGCTG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 33400
CGTCGGGGCGGACAAGCGCCAGCTCCAGCGCGACAAGGCGGAGGAACCTCAGGACCCACACGCGGGGCTGTACGACCGGCCCGTGAAGGTAGCCGCTCGAC
Q P A L F A V E V A L F R L L E S W G V R P D M L A **G H S** I G E L
ATa

GCGGGGCCACGCTTCGGGTGTGGTTCGCTGCCGATGCCGCAAGATCGTGGCCGCGGGGGCGCGCTGATGCAGGGCTGCCCAAGGGCGGGCGGA
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 33500
CGCCGCGGGTGCAGAGCCACACACAGCGACGGCCTACGGCGTTCTAGCACCGCGCGCCCCGCGGACTACGTCCGCGACGGGTTCGCCGCGCGT
A A A H V S G V W S L P D A A K I V A A R G A L M Q A L P K G G A M

TGGTCGCGATCCAGGCGTCCGAGGAAGAGGTTCGCCCTCCGACGACCTGCCGGAGACCGTGGGCATCGCTGCGGTCAACGGCCCGTCCCGTTGTCATTT
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 33600
ACCAGCGTAGGTCCGAGGCTCCTTCTCCAGCGGAGGCTGCTGGACGGCCTTGGCACCCGTAGCGACGCCAGTTGCCGGGCGTAGGCAACAGTAAAG
V A I Q A S E E E V A S D D L P E T V G I A A V N G P S S V V I S

GGGTGTGGCCCGGACGTCGATGCGGTTCGCGGAGCGCTGGCGTGGCGGGCCGCAAGGTACGCGGCTGAAGGTACGCCACGCGTTCCTACTGCGCGTTG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 33700
CCCACACCGCGGCTGCAGCTACGCCAGCGCCTCGGACCGCACGCGCCCCGGCGTTCCAGTGCGCCACTTCCAGTCCGTTGCGCAAGGTGAGCGGCAAC
G V A A D V D A V A E R W R A A G R K V T R L K V S H A F H S P L

ATGGACCCGATGCTGGACGAGTTCCTGACGGTTCCTTGAGGGCGTGTCTTCGAGGCGCCGAGATTCCGATCGTGTGACGCTGACCGCGCGCAGGCGT
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 33800
TACCTGGGCTACGACCTGCTCAAGGCATGCCAGGAACCTCCCGCACAGGAAGCTCCCGGGCTCTAAGGCTAGCACAGCTGCGACTGGCCGCGCGTCCGCA
M D P M L D E F R T V L E G V S F E A P Q I P I V S T L T G A Q A S

CCGCCGAGGAGCTGGCCTCCCGGAGTACTGGGTGGGGACGCTGCGGAGTCCGTGCGCTTCCACGACGCCGTCGTGGCCCTGCGCGAGCAGGAAGTCGA
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 33900
GGCGGCTCCTCGACCGGAGGGGCTCATGACCCACGCGTGCACGCGCTCAGGCACGGAAGGTGCTGCGGCAGCACCGGGACGCGCTCGTCTTACGCT
A E E L A S P E Y W V R H V R E S V R F H D A V V A L R E Q E V D

CGTGTTCGTGGAGGTCCGCCCCGGGGGACGCTGTCCGCTCTCGGGCTGGCAGCGCGCCGACCGGTGTTCCTGCCCGCGCTGCGGACGGACAGGCC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 34000
GCACAAGCACCTCCAGCCGGGCCCGCTGCGACAGGCGAGAGCCGAGCCGTGCGCGGGCTGCGCCACAAGGACGGGCGCGACGCTGCGTGTCCGGG
V F V E V G P G G T L S A L G V G S A P D A V F L P A L R T D R P

GAGGAGATCGCGCTCGTCACGACGGTGGCCGCGCTGTACACGACGGCATGCCGGTTCGACTGGACCGCCCCGCTTCGACGGCAGCGGCCCGCCGAGGT
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 34100
CTCCTTAGCGGAGCAGTGTGCCACCGGCGGACATGTGCGTCCCGTACGGCCAGCTGACCTGGCGGGCGAAGCTGCCGTGCCCGGGGCGGCTCCA
E E I A L V T T V A A L Y T Q G M P V D W T A P L R R Q R R P G R S

Anexo I

CGACCTGCCGACGTACGCGTTCCAGCGCGAGCGCTTCTGGCTGGACGTGCCCGCCGACGCGCGCCGAGCGGATGTCGCCGGTGGAGGCGCGTTCTGG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 34200
GCTGGACGGCTGCATGCGCAAGGTCGCGCTCGCAAGACCGACCTGCACGGGCGGCTGCGCGCGGGCTCGCTACAGCGGCCACCTCCGCGCAAGACC
T C R R T R S S A S A S G W T C P P T R A P S G C R R W R R G S G

GAGGCCATCGCGGAGGGCGACCTCGCCGACCTGGCCCGCATGCTGAACGTGTCGTGCGGACACCCGCTGAGCGACGTGCTGCCCATGCTGTCCGCTGG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 34300
CTCCGGTAGCGCTCCCGCTGGAGCGGCTGGACCGGGGCGTACGACTTGCACAGCAGCCTGCTGGGCGACTCGCTGCACGACGGGTACGACAGGGCGGACC
R P S R R A T S P T W P R M L N V S S D D P L S D V L P M L S A W

CGCACGGAGCAGCAGGGCCGGCCACCGTGGACGGCTGGCGCTACCGGATCGACTGGAGCCCGATAGCCGCCGCGAGCCGTCAACTGGACGGCGCCTGGC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 34400
GCGTGCCTCGTCTCCCGGCCGGTGGCACCTGCCGACCGGATGGCTAGCTGACCTCGGGCTATCGGCGGGCTCGGCAGTTGACCTGCCGCGGACCG
R T E Q Q G R A T V D G W R Y R I D W S P I A A A S R Q L D G A W L

TGGTCTGGCGCCGAGCGGAGACCCGCGCGGAGTGGGTGCGTGGCGCTGGCCCGCAACGGTGGGACGTCCACGTCCACGTCTGGACGTGGACCC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 34500
ACCAGCACCGCGGCTCGCTCTCGGCGCGCCGCTACCCACGCACCGCGGACCGGGCGTTGCCACGCCTGCAGGTGCAGGTGCAGGACCTGCACCTGGG
V V A P S G E P R G E W V R A A L A R N G A D V H V H V L D V D P

GGCGACGGTTCAGTGGCGGAGCGGCTGAGCAGCTGCCCGCTCTCGCGGAGTGGTCTCGCTGCTGGGGCTCGCCGACGCGGGCGAGTCGGCGGTGCC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 34600
CCGCTGCCAGCTCACCCGCTCGCGACTCGCTCGACGGGCGAGAGCCGCTACCCAGAGCGACGCCCGGAGCGGCTGCGCCGCTCAGCCGCCACGGG
A T V E W A E R L S E L P A L G G V V S L L G L A D A G E S A V P

GCAGGGGTTCAGGCCACGATCGGTCTGCTGCGGGCGCTCGGCGACGCGGGTTCGAGGCGCCGTTGTGGTGCCTGACGTCCGGTGGCGGGCGGCGG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 34700
CGTCCCAGTGCCTGCTAGCCAGACGACGCCCGGAGCCGCTGCGCCACAGCTCCGCGCAACACCACGGACTGCAGGCCACGCCACCGCCCGCGG
A G V T A T I G L L R A L G D A G V E A P L W C L T S G A V A A A D

ACGGCGACGCGGTTCGCGGGTTCGAGCAGTCTGGTGTGGGGTCTGGGCCGGGTGGCCGCGCTGGAGCAGCCGGGCCGCTGGGGCGGTCTGGTGGACCT
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 34800
TGCCGCTGCGCCAGCGGCCAAGCTCGTCAACACACCCAGACCCGCGCCACCGGCGGACCTCGTCCGCCCGGCGACCCCGCCAGACCACCTGGA
G D A V A G F E Q S L V W G L G R V A A L E Q P G R W G G L V D L

GCCCGCGGCGCCGACGAGACGACCGGTGACCTGCTGGCCGCGGTGCTCGCCGGTACGGGCGGCGAGGACCAGGTGCGCTGCGCGGCAACGGATCC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 34900
CGGGCGCGCGGGCTGCTCTGCTGGCCACTGGACACCGGCGCCACGAGCGCCATGCCCGCGCTCCTGGTCCAGCGGACGCGCCGTTGCCCTAGGAG
P A A P D E T T G D L L A A V L A G T G G E D Q V A L R G N G I L

GGCCGAGGCTGGTGCCTGCGCGCCGCTGGGCGGTGCCGCTCCGAGGCTGGTGCCTGCGCGCGGAAACGGTCCCTGGTGACCGGCGGACCGGCGCGCTCGGCG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 35000
CCGGCTCCGACCACGGGCGGCGGACCCGCGCAGGCTCCCGACCGGCGCGCCTTGCCAGGACCACTGGCCCGCTGGCCGCGCGAGCCGCG
G R R L V P A P L G G A A S E G W S P R G T V L V T G **G T G A L G A**
KR

CGCACGTGCCCGGTGGCTGGTGGCGAAGGGCGCCGAGCACCTGCTCCTCACAGCAGGCGCGGTGCCGCGGCGAAGGCGCGGAGGAACCTGCGTGGCGA
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 35100
GCGTGCAGCGGGCCACCGACCACCGCTTCCCGGGCTCGTGGACGAGGAGTGGTCTCCGCGCCACGGCGCGGCTTCCGCGGCTCCTGACGCACGCCT
H V A R W L V A K G A E H L L L T S R R G A A A E G A E E L R A E

GCTGACCAGCTCGGCGCGCGCTCACCATCGCGGCTGCGACGCCCGGACCGGGACCGCTGGCCGACGTACTCGCTCCGTGCCCGCGGAGTTCCCG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 35200
CGACTGGTGCAGCCGCGCGCAGTGGTAGCGCCGACGCTGCGGCGGCTGGCCCTGCGCGACCGGTGCATGAGCGGAGGACGGGCGGCTCAAGGGC
L T T L G A R V T I A A C D A A D R D A L A D V L A S V P A E F P

Anexo I

CTCACGCGGTTCGTGCACGCGGCCGGCATCGAGCGCACGGCTCCGCTCGCGGAGCTGGACGAGGACGACCTGGCCGGTGTGCTCGCGGCGAAGTTCGACG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 35300
GAGTGGCGCCAGCACGTGCGCCGGCCGTAGCTCGCGTGCCGAGGCGAGCGCCTCGACCTGCTCCTGCTGGACCGGCCACACGAGCGCCGCTTCCAGCTGC
L T A V V H A A G I E R T A P L A E L D E D D L A G V L A A K V D G
SacI
GCGCGGGAACCTGCACGAGCTC
-----+-----+---- 35323
CGCGCGCCTTGGACGTGCTCGAG
A R N L H E L

ANEXO II

Secuencia de nucleótidos de 13955 pb que alberga la región 3' terminal del gen *flvC* y el extremo 5' terminal del gen *flvD*. Se muestra la secuencia de aminoácidos deducida de cada región abierta de lectura. En negrita se marca la secuencia de nucleótidos codificante. Y con flecha discontinua se señala el inicio de transcripción del gen *flvD*.

CCC GGATCGGCCGAGGCGGCCGCGCGACGACGGCTGCTGGACAGTTGTACGTGCGCGGTGTCAGGTGCGGTGGGAGGCGTTCTTCGCAGGGCGCGGC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 100
GGGCCTAGCCGGCCTCCGCCGGCGCGCTGCTGCCGACGACCTGTCAACATGCACGCCACAGGTCACGCGCACCCCTCCGCAAGAAGCGTCCCGCGCCG
P D R P E A A A A T T A A G Q L Y V R G V Q V A W E A F F A G R G
.....▶ *flvC*
GCACAGCGCGTGGACCTGCCGACGTACGCGTTCCAGCATGAGCGCTACTGGCCCAGGCACTGGCGCTGACCGCGACGTCACGGCGGCCGGCTCGGTT
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 200
CGTGTGCGGCACCTGGACGGCTGCATGCGCAAGGTCGTACTCGCGATGACCGGGCTCCGTGACCGCGACTGGCCGCTGCAGTGCAGCCGGCCCGGAGCCAA
A Q R V D L P T Y A F Q H E R Y W P E A L A L T G D V T A A G L G S

CGGCGGAGCATCCGTTGTTGGGTGCGGCTGTTGCTGTGGGTGGCACGGATGGTGTGCTGCTGACTGGTCTTTTCGGTTCCAGACGCATCCGTGGCTGGC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 300
GCCGCTCGTAGGCAACAACCCACGCCGACAACGACACCCACCGTGCCTACCACACGACGACTGACCAGCAGAAAGCCAAGTCTGCGTAGGCACCGACCG
A E H P L L G A A V A V G G T D G V L L T G R L S V Q T H P W L A

TGAGTATGTGGTTTCAGGTTTCGGCGCTGCTGCCTGGTGGCGCTTCGTGAGCTGGCGATCCGCGCGGGGGACCAGGTTCGGGTGCGATCTGGTGGAGGAG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 400
ACTCATAACCAAGTCCCAAGCCGCGACGACGGACCACGCCGAAGCAGCTGACCGCTAGGCGCGCCCCCTGGTCCAGCCCACGCTAGACCACCTCCTC
E Y V V Q G S A L L P G A A F V E L A I R A G D Q V G C D L V E E

TTGACGCTTGAGGCGCCGCTTGTCTGCCTGCGTCTGGGTGCGGTGCGGGTGCAGTTGTGGTTGGTTCGGTTCGGATGGGGCGGCAAGCGTGAGCTGACGT
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 500
AACTGCGAACTCCGCGGCGAACAGGACGGACGCGACGCCACGCCACGCTCAACACCCAACCACGCAGCCTACCCCGCCGTTGCGACTCGACTGCA
L T L E A P L V L P A S G S V R V Q L W V G A S D G A G K R E L T F

TCTATTCCGGTTCGGGTGATGTTGAGGACGGGCGTTCTTGGATCCGGCATGCCGTGGGTGTGCTGCGCGAGGGTGGCCGTTCTGAGGGGGTTTCCCTGGA
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 600
AGATAAGGCCACGCCACTACAACCTGCCCCGAAGGACCTAGGCCGTACGGCACCCACAGCAGCGCTCCACCGGCAAGACTCCCCAAAGGGACCT
Y S G A G D V E D G R S W I R H A V G V L R E G G R S E G V S L E

GGTGTGGCCTCCGGCTGGGCTGAGGTTGTGGATCTGGAGGGGTTCTACGACGAGGGGCTGACGCCGCTTCGCCCTACGGGCCGTTTCCAGGGTTG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 700
CCACACCGGAGGCCGACCCGACTCCAACACCTAGACCTCCCAAGATGCTGCTCCCCGACTGCGGCCAGAGCGGATGCCCGGCCACAAGTCCCAAA
V W P P A G A E V V D L E G F Y D E G A D A G L A Y G P V F Q G L

CGGCGCGCTGGCGCAAGGACGGCGAGGTTTCGCCGAGGTGCGGTTGCCGATGGTGTGAAGTTCGGACGGCTTCGGTCTGCATCCGGCGCTGCTTGACG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 800
GCCCGGCGCACCGCTTCCTGCGCTCCACAAGCGGCTCCAGCGCAACGGGCTACCACACTTCAGCCTGCCGAAGCCAGACGTAGGCCGCGACGAACTGC
R A A W R K D G E V F A E V A L P D G V K S D G F G L H P A L L D A

CGACGTTGCACACGGTTCGGTCTCGCGGACAGCGGGACGTCGACGGCAAGGTGCCGTTTCGGTGGTTCGGGTGTGCGGCTGCACGCCCTCCGGCGGACGGC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 900
GCTGCAACGTGTGCCAGCCAGAGCGCTGTGCGCCTGCAGCTGCCGTTCCACGGCAAGCGCACCAGCCACACGCCGACGTCGGGAGGCCGCGCTGCCG
T L H T V G L A D S A D V D G K V P F A W S G V R L H A S G A T A

ACTGCGGCTGCGCCTCGCCCTGCGGAGCGGACGCGGTGACGCTGACGCTCGCCGACGGCGGGGCGCCCGTTCGGACGGTGGACTCCCTCGTCTCTG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 1000
TGACGCGCACGGGAGCGGGGACGGCCTCGCTGCGCCACTGCGACTGCGAGCGGCTGCCGCGCCCGGGGCGAGCGCTGCCACCTGAGGGAGCAGGAC
L R V R L A P A G A D A V T L T L A D G A G A P V A T V D S L V L

Anexo II

CGTCCCCTAGCGGAGACGCTGTCGAGGCAGGACGGGGAATCGCTGTTTCGGCGTTCGACTGGGTGTCCGTGCCCTGTCCGACGGCGACGCCCCGACGGTCCG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 1100
GCAGGGCATCGCCTCTGCGACAGCTCCGTCTGCCCTTAGCGACAAGCCGACAGTACCCACAGGCACGGGGACAGGCTGCCGCTGCGGGGCTGCCAGC
R P V A E T L S R Q D G E S L F G V D W V S V P L S D G D A P T V A

CATGGGCCGACCTGGAAGCACTGACCGAGCCGAGTCGGCGGAAGTGCCTGGGCTTCGTGGTCTGTCTGCCCCGACACCTCGGGCTCCGACCCGGCCGT
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 1200
GTACCCGGCTGGACCTTCGTGACTGGCTCGGCGTCAGCCGCCTTGACGGCCCAAGCACCAGGACAGCAGCAGGGGCTGTGGAGCCGACAGGCTGGGCCGCA
W A D L E A L T E P Q S A E L P G F V V L S C P D T S A S D P A V

CGGCGCGCACGACGCCCGCACTGGGCGTGAACGCGGTCCAGAGCTGGCTGACCGACGAGCGCTTCGACGCGGCCCGCTGGCCGTCTGACAGGGGC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 1300
GCCGCGCTGCTGCGGCGGTGACCCGCGACTTGCCAGGTCTCGACCGACTGGCTGCTCGGAAGCTGCGCCGGGCGGACCGGCAGCACTGGTCCCCG
G A H D A A H W A L N A V Q S W L T D E R F D A A R L A V V T R G

GCGATCGCCGCTCGGCCGATGAGGCGGTACGGGACGTCGCGCAGGCCGCGTGTGGGGCTGGTGCCTCCGCCAGTTCGGAGAACCCGGACCCGGATCG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 1400
CGCTAGCGGGGAGCCGGCTACTCCGCCATGCCCTGCAGCGGTCCGGCGGCACACCCCGGACCAGCCAGGCGGGTACGCTCTTGGCCCTGGCCTAGC
A I A A S A D E A V R D V A Q A A V W G L V R S A Q S E N P D R I V

TCCTGGTTCGACTTCGACGAGGAGACGGCCTCGCTGGACGCGCTGCCGTCCGCGCTGGCCACCGGTGAGCCGAGTTGGCGCTCCGCTCGGGCACCGCGAG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 1500
AGGACCAGTGGAGCTGCTCCTCTGCCGGAGCGACCTGCGCGACGGCAGGCGCGACCCGGTGGCCACTCGGCGTCAACCGCGAGGCGAGCCCGTGGCGCTC
L V D L D E E T A S L D A L P S A L A T G E P Q L A L R S G T A S

CGCCCCGCGCTGGCCAGGCTCAGCCGCTCCGTGGTGGACGACCCCGGCTTCGGGCGGGTGCCTGCTCATCACGGGCGGACGGGCACGCTGGGCGGC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 1600
GCGGGGCGGGACCCGTCAGTCGGGAGGACACCTGCTGGGCGCAAGCCCGGCCACGGCAGAGTAGTCCCGCGCTGCCCGTGGACCCCGCC
A P R L A R V S P S V V D D P G F G P G A V L I T G A T G T L G G

CTCATGGCCCGCACCTGGTTCGCGGAGCGCGGCTACGCAACCTGCTGCTGGTGGAGCGGCGGCGCCCGCCGGAAGGCGCCGAGGAACCTGCGGCGCG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 1700
GAGTACCGGGCCGTGGACAGCGCCTCGCGCCGATGCGTTGGACGACGACCACTCGTCCGCGCCGGCGGGCGGCTTCCGCGGCTCCTTGACGCGCGGC
L M A R H L V A E R G V R N L L L V S R R G A A A E G A E E L R A E

AACCTCGCCGCGCAGGGCGCCGAAGTAGTGCCTCGCCGCTGCGATGTGGCCGACCGCGAGGCTGTTCGCGCGCTGCTGGCCGCGCACGACATCTCGGCGCT
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 1800
TTGAGCGGCGCTCCCGCGGCTTCATCACGAGCGGCGGACGCTACACCGGCTGGCGCTCCGACAGCGGCGGACGACCCGGCGGTGCTGTAGAGCCGCA
L A A Q G A E V V L A A C D V A D R E A V A A L L A A H D I S A V

CGTCCACACCGCGGTGTCTTCGACGACGGCGTGATCAGCTCGCTCACCCCGGAGCGGATCGACACCGTCTTCGCCCGAAGGTCGACGCGGCTGGCAC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 1900
GCAGGTGTGGCGGCCACAGGAGCTGCTGCCGCACTAGTCGAGCGAGTGGGCGCTCGCTAGCTGTGGCAGAAGGCGGGCTTCCAGTGCGCCGGACCGTG
V H T A G V L D D G V I S S L T P E R I D T V F R P K V D A A W H

CTGCACGAGCTGACCCGCGACCTGAACCTGTCCGCGTTCGTCTCTCTCGGTCTCGGGCACGCTGGGCGCCCCCGGCCAGGGCAACTACGCGGCGG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 2000
GACGTGCTCGACTGGGCGTGGACTTGGACAGGCGCAAGCAGGACAAGAGGAGCCAGAGCCCGTGCACCCGCGGGGCGGTTCCCGTTGATGCGCCGCC
L H E L T R D L N L S A F V L F S S V S G T L G A P G Q G N Y A A A

Anexo II

CGAACGCGTTCTCGGACGCGCTCGCGCAGCACCGCCGTGCCGCCGACTGCCGCCACCTCGCTCGCCTGGGGCCTGTGGGCCGAGGCCAGCGGCATGAC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 2100
GCTTGCACAAGGACCTGCGCGAGCGCTCGTGGCGGCACGGCGGCCTGACGGGCGGTGGAGCGAGCGGACCCCGGACACCCGGCTCCGGTTCGCCGTACTG
N A F L D A L A Q H R R A A G L P A T S L A W G L W A E A S G M T

CGGGGACCTCGGCGAGGCCGACGTGAAGCGGATGGGCCGGGGCGCCTGCTCCCGCTCCGCACCGCCGAGGGCCTCGCCCTGTTTCGACGCGGGCGGGCTCG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 2200
GCCCCTGGAGCCGCTCCGGTGCACCTTCGCTACCCGGCCCCGCCGACGAGGGCGAGGCGTGGCGGCTCCCGGAGCGGGACAAGCTGCGCCGCCGAGC
G D L G E A D V K R M G R G G L L P L R T A E G L A L F D A A G S

ACCGGCCGCGCCGCGCTCGTACCGTGCCTTGACCTGGCCGTGCTGCGCCGAGGCCCGCACCCGCCCGCTCCCGCACCTCCTGCGCGGCTGGTGC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 2300
TGGCCGCGCGCGCGAGCATGGCCACGGGACCTGGACCGGCACGACGCGGCGCTCCGGGCGTGGCGGGGCGAGGCGTGGAGGACGCGCCGGACCAGC
T G R A A L V P V P L D L A V L R R E A R T R P V P H L L R G L V R

GCGGTACCGCGCGCAGGACCGCCGAGTCCGGCCCCGGCGGAGTTCGGCGCTGGCCCGCTCGCTCGCCGGTATGAGCCTCGCGGACCGGGAGAAGGCCCT
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 2400
CGCCATGGCGCGCTCCTGGCGGCTCAGGCCGGGGCCGCGCTCAAGCCGCGACCGGGCGAGCGAGCGGCCATACTCGGAGCGCCTGGCCCTCTTCCGGGA
G T A R R T A E S G P G A S S A L A R S L A G M S L A D R E K A L

CCTCGACCTCGTACTCGGTACGTCGGCCGTGGTGCFCGGCCACAGCTCGGCGGGCGCCATCGCACCCGACCGCGCTTCAAGGACCTCGGCTTCGACTCG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 2500
GGAGCTGGAGCATGAGCCAGTGCACCGGCACCACGAGCCGGTGTGAGCCGCCCGCGGTAGCGTGGGCTGGCGCGCAAGTTCCTGGAGCCGAAGCTGAGC
L D L V L G H V A V V L G H S S A G A I A P D R A F K D **L G F D S**
ACP

CTCAGCTCCGTGAACTGCGTAACCACCTGGGCGCGGGACCGAACTGGCCCTGCCCGCCACGCTCGTCTTCGACTACCCGACCCCGCCGCACTCGCCG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 2600
GAGTCGAGGCGCTTGCAGCATTGGTGGACCCGCGCGCTGGCTTACCGGGACGGCGGTGCGAGCAGAAGCTGATGGGCTGGGGCGCGCTGAGCGGC
L S S V E L R N H L G A A T E L A L P A T L V F D Y P T P A A L A A

CCCACCTCGGCGCCGAACTCGCGGGCGCCACCGTGGTTCGAAACGGTACCGGCACGAGCGTCCGATGGCCGACGAACCGATCGCGATCGTCCGATGAG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 2700
GGGTGGAGCCGCGGCTTGCAGCCCCGCGTGGCACCAGCTTTCAGTGGCCGTGCTGCCAGCGGTACCGGCTGCTTGGCTAGCGCTAGCAGCCGTACTC
H L G A E L A G A T V V E T V T G T T V A M A D E P I A I V G M S

CTGCCGCTTCCCGCGCGGCTCGAGTCGCCCGAGGACCTGTGGCGCTCGTCGCGTCCGGCGCGGACCGCGTGTCCGGCTTCCCGACCGATCGCGGCTGG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 2800
GACGGCGAAGGGGCCCGCGAGCTCAGCGGGCTCCTGGACACCGCCGAGCAGCGCAGGCGCGCCTGCGGCGACAGGCCGAAGGGCTGGTAGCGCCGACC
C R F P G G V E S P E D L W R L V A S G A D A L S G F P T D R G W

GACGTGGAGTCGCTCTACGACCCGACCCGACCACGCGGGCAGACCTACGCCGTGAGGGCGGCTTCTGACGACGTCGCGGAGTTCGACGCGGACT
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 2900
CTGCACCTCAGCGAGATGCTGGGCTGGGCTGGTGCGCCGTGCTGGATGCGGGCACTCCCGCGAAGGACGTGCTGCAGCCGCTCAAGCTGCGCCTGA
D V E S L Y D P D P D H A G T T Y A R E G G F L H D V G E F D A D F

TCTTCGGCATCTCGCCGCGGAGGCCCTGGCCATGGACCCGACGAGCGACTGCTCCTCGAAACCTCCTGGGAGGCGTTCGAGCGGGCGGGCATCGACCC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 3000
AGAAGCCGTAGAGCGCGCGCTCCGGGACCGGTACCTGGCGTCTGCTGACGAGGAGCTTGGAGGACCTCCGCAAGCTCGCCGCCGCTAGCTGGG
F G I S P R E A L A M D P Q Q R L L L E T S W E A F E R A G I D P

GCACGCGGTACGCGGACCGCTCCGGTGTGTTTCGTCGGCACCAACGGCCAGGATTACGCGCGCTGTTGATGGCCGCCAGGAGAGCGCCGAGGGCTAC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 3100
CGTGCCTATGCGCCGTCGGCGAGGCCACACAAGCAGCCGTGGTTCGCCGTCTAATGCGCCGCAACTACCGCGGGTCTCTCGCGGCTCCCGATG
H A V R G S R S G V F V G T N G Q D Y A A L L M A A Q E S A E G Y

Anexo II

GCGGGCACGGGCAGCGCGGCGAGCGTCATCTCGGGCCGCTCTCGTACACGTTTCGGCCTCGAAGGCCCGGCCGTACCGTGGACACGGCGTGCTCCTCGT
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 3200
CGCCCGTGCCCGTCGCGCCGCTCGCAGTAGAGCCCGGCGGAGAGCATGTGCAAGCCGGAGCTTCGGGGCCGCGAGTGGCACCTGTGCCGCACGAGGAGCA
A G T G S A A S V I S G R L S Y T F G L E **G P A V T V D T A C S S S**
K S

CGCTGGTTCGCCCTGCACCTCGCGGGCCAGGCGCTGAACAAGGGCGAGTGGCAGCTGGCCCTGGTTCGGCGGTGTGACGGTGATGTCGACGCCGGGTGCGTT
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 3300
GCGACCAGCGGGACGTGGAGCGCCCGGTCCGCGACTTGTTCGCCGCTCAGCTCGACCGGGACCAGCCGCCACACTGCCACTACAGCTGCGGCCACGCAA
L V A L H L A G Q A L N K G E C E L A L V G G V T V M S T P G A F

CATCGAGTTCAGCCGCCAGCGCGGCTCGCGCCGAGCGGCCGCTGCAAGGCGTTTCGCCGAGGGCCGCGACGGCACCGGCTGGGGCGAGGGCGTGGGCATG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 3400
GTAGCTCAAGTCGGCGGTTCGCGCCGAGCGCGGCTGCGCGCAGCTTCGCCAAGCGGCTCCCGCGCTGCGTGGCCGACCCCGCTCCCGCACCCGTAC
I E F S R Q R G L A P D G R C K A F A E G A D G T G W G E G V G M

CTGCTCGTGGAGCGCCTTTCCGACGCCCGCGCAACGGCCACGAGGTGCTTTCGGTTCGTCGGCGCTCCGCGGTGAACCAGGACGGCGCGTGAACGGCC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 3500
GACGAGCACCTCGCGAAAGGCTGCGGGCCGCTTCCCGTGTCCACGAACGCCAGCAGGCGCCGAGGGCGCCACTTGGTCTGCGCGCAGCTTCCCG
L L V E R L S D A R R N G H E V L A V V R G S A V N Q D G A S N G L

TGACGGCCCCGAACGGCCGTCGACGAGCGGTCATCCGCCAGGCACTGGCCAGCGCGGGCTGGCCGCGCCGACGTGGACGCGGTGAGGCGCACGG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 3600
ACTGCCGGGCTTCCCGGCGAGCGTCGTCGCGCAGTAGGCGGTCCGTGACCGGTTCGCGCCCGGACCGGGCGCCGGCTGCACCTGCGCCAGCTCCGCGTGC
T A P N G P S Q Q R V I R Q A L A S A G L A A A D V D A V E A H G

CACCGGCACGTTCGCTGGGGACCCGATCGAGGCGCAGGCCCTGCTCGCCACGTACGGGCAGGACCGCGGGGTGACCCGCTGTGGCTCGGCTCGGTCAAG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 3700
GTGGCCGTGCAGCGACCCGCTGGGCTAGCTCCCGTCCGGGACGAGCGGTGCATGCCGCTCCTGGCGCCGCCACTGGGCGACACCGAGCCGAGCCAGTTC
T G T S L G D P I E A Q A L L A T Y G Q D R G G D P L W L G S V K

TCGAACGTTCGGTTCACACGCAGGCCGCGGCCGGTGTTCGGGTTGATCAAGATGGTTCATGGCGATGCGGGAGGGCGTGTGCCGCGCACCTGCACGTGC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 3800
AGCTTGCAGCCAGTGTGCGTCCGGCGCCGCCACAGCGCCACACTAGTTCTACAGTACCCTACGCCCTCCCGCACGACGGCGGTGGGACGTGCAGC
S N V G H T Q A A A G V A G V I K M V M A M R E G V L P R T L H V D

ACGAGCCCTCGACGCACGTGACTGGTCCGCGGGCGCGTCCGGCTCCTGACGGAGACCCAGCCGTGGCCCGAGACGGGTGCGCCGTGGCGGGCGGGTGT
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 3900
TGCTCGGGAGCTGCGTGCAGCTGACCAGGCGCCCGCGGCGAGGCGGACTGCTTGGTTCGGCACCGGGCTTGGCCAGCCGGCACCCCGCCGCCCCA
E P S T H V D W S A G A V R L L T E T Q P W P E T G R P W R A G V

GTCTGCTTCGGCGTGAGCGGCACGAACCGGCACGCCATCATCGAACAGGCCCGCGCTCCAGGAGTTGGACGTCGACGGCGCCGGAGCCCGACGGG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 4000
CAGCAGCAAGCCGCACTCGCCGTGCTTGGCGTGCAGTAGTGTTCGGGGCCGGCAGGTCTCAACCTGCAGCTGCGCCGCGGCTCGGGCTGCC
S S F G V S G T N A H A I I E Q A P A V Q E L D V D A A P E P D G

CTCGCGCGTGGGTTCCTTTCGGCCCGGAGCGCCGAGGCGCTGCGGGAGCAGGCCAGGCGTCTGAAGTCTTCGTTGGACGCGACCGGTGAGCTGGCCGTGT
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 4100
GAGCGCGCACCCAGGAAAGCCGGGCTCGCGGCTCCCGCAGCCCTCGTCCGGTCCGACACTTCAGGAAGCACCTGCGTTCGCGACTCGACCGGCACA
L A P W V L S A R S A E A L R E Q A R R L K S F V D A T R E L A V S

Anexo II

CGGACGTGGGTCTGTCGCTGGTGTCTGTCGCGGGCGGTGCTTGAGCACCGTGCGGTGGTCTGCGCCGGTGACCGGGACGAGTTCTGGCCGCTCTGGCGGC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 4200
GCCTGCACCCAGCACGCGACCACAGCAGCGCCCGCCACGAACTCGTGGCACGCCACCAGCAGCGGCCACTGGCCCTGCTCAAGGACCGGCGAGACCGCCG
D V G R A L V S S R A V L E H R A V V V A G D R D E F L A A L A A

GGTGGCCGACGGCCAGGCGAGCGGGCGCGTGGTCTCGGGCGTCCGGATGAGCCCGCGGGTGGGTTTCTGTTCTCGGGTCAGGGTTCGACGCGCCTT
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 4300
CCACCGGCTGCCGGTCCGCTCGCCGCGGCACCAGAGCCCGCAGCGCTACTCGGGCCGGCCACCCAAAGGACAAGAGCCAGTCCCAAGCGTCCGCGGAA
V A D G Q A S G A V V S G V A D E P G R V G F L F S G Q G S Q R L

GGTATGGGGCGTGAAGTGTCCGAGCGTTTCCCGGTGTTCCGCGGGCGCTGGACGAGGTCTTGGGAGTTCGACCCGCTCTGTGCGTGAGGTGTTGTTTCG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 4400
CCATACCCCGCACTTGACAGGCTCGCAAAGGGCCACAAGCGCCGCGGACTGCTCCAGGAACGCCTCAAGCTGGGCGAGACACGCACTCCACAACAAGC
G M G R E L S E R F P V F A A A L D E V L A E F D P S V R E V L F G

GTCAGGACCGGACGCCCTGAATGAGACCGGGGTGACGCAGCCCGCGCTGTTCCGCGGTTCGAGGTGGCGCTGTTCCGCTCTCTGAGTCTGGGGCGTCCG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 4500
CAGTCTGCGCTGCGGGACTTACTCTGGCCCACTGCGTGGGCGGACAAGCGCCAGTCCACCGCGACAAGCGGAGGAACTCAGGACCCCGCAGGC
Q D A D A L N E T G V T Q P A L F A V E V A L F R L L E S W G V R

CCCCGACATGCTGGCGGGCCACTCGATCGGTGAGCTGGCCCGCGCATGTGTCGGGTGTTGGTTCGCTGCCGGATGCCGTAAGGTGGTGTCCGCGCGG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 4600
GGGGCTGTACGACCGCCCGGTGAGCTAGCCACTCGACCGGCGGCGGTACACAGCCACAGACCAGCGACGGCCCTACGGCAGTTCACCACAGGCGCGCC
P D M L A **G H S** I G E L A A A H V S G V W S L P D A V K V V S A R
ATp

GGTGCCTGATGCAGGCGCTCCCGAGGGCGGGCGATGGTTCGGATCCAGGCGACGGAGGCGAGGTTCGCGGCCGACGAAGTGCAGGACACCGTCCGACA
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 4700
CCACGCGACTACGTCGCGAGGGGCTCCCGCCGCGTACCAGCGCTAGGTCCGCTGCCTCCGGCTCCAGCGCGGCTGCTTGACGGCTGTGGCAGCTGT
G A L M Q A L P E G G A M V A I Q A T E A E V A A D E L P D T V D I

TCGCGGCCGTGAACGGCCCCGCTCGGTGGTCTGTTTCCGGTGTGGCCGCTGATGTCGGTGCGGTTCGCGGAGCGGTGGCGTGAGCGGGTTCGCAAGGCGAC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 4800
AGCGCCGGCACTTGGCGGGGCGGAGCCACCAGCAAAGCCACACCGGCGACTACAGCCACGCCAGCGCCTCGCCACCGCACTCCGCCAGCGTTCCGCTG
A A V N G P A S V V V S G V A A D V G A V A E R W R E A G R K A T

CCGGCTGAAGGTACGCCACGCGTTCCTACTCGCCGCTGATGGACCCGATGCTGGACGAGTTCAAGCGGGTCTCGCGGGCGTGTGTCAGAGGCGCCCGAG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 4900
GGCCGACTTCCAGTCCGTCGCAAGGTGAGCGGCGACTACCTGGGCTACGACCTGCTCAAGTTCGCCAGGAGCGCCCGCACAGCATGCTCCGCGGGGTC
R L K V S H A F H S P L M D P M L D E F K R V L A G V S Y E A P Q

ATCCCGATCGTGTGACACTCACCGGCCGACCGGCCACCGCCGAGGAGTGGCTCCCGGACTACTGGGTGCGGCACGTGCGGAAGGCGGTGCGCTTCG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 5000
TAGGGCTAGCACAGCTGTGAGTGGCCGGCTGGCCGGTGGCGGCTCCTCGACCGGAGGGGCTGATGACCCACGCCGTGCACGCTTCCGCCACCGGAAGC
I P I V S T L T G R P A T A E E L A S P D Y W V R H V R K A V R F A

CAGACCCGTTGAGCGCCCTCGTGGAGGAGGCGTGGCGACGTTTCGTCGAGGTTCGGTCCCGCGGGCGCTGTCGGCGCTGGGGCAGGAGTCCGCGCCGGA
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 5100
GTCTGCGGCACTCGCGGGAGCACCTCCTCCCGCACCGCTGCAAGCAGCTCCAGCCAGGGCCGCCCGCGACAGCCGCGACCCCGTCTCAGGCGCGGCT
D A V S A L V E E G V A T F V E V G P G G A L S A L G Q E S A P D

Anexo II

CGCGGCCTTCGTACCCGTGCTGCGGAGCGATCGGCCGGAGGCCGTGGCCGCCACGACGGCTGCCGGGCACCTGCACGTGCGCGGTGTCCAGGTCGACTGG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 5200
GCGCCGAAGCATGGGCACGACGCCTCGCTAGCCGGCCTCCGGCACCGCGGTGCTGCCGACGGCCCGTGGACGTGCACGCGCCACAGGTCCAGCTGACC
A A F V P V L R S D R P E A V A A T T A A G H L H V R G V Q V D W

CAGCCCTTCTTCGACGGCAGCGGTACGCGACGGTGGACCTGCCGACGTACGCGTTCAGCGCGAGCGCTACTGGCTGGACGCCACATACACGACCGGTG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 5300
GTCGGGAAGAAGCTGCCGTGCGCATGCGCTGCCACCTGGACGGCTGCATGCGCAAGGTGCGCTGCGGATGACCGACCTGCGGTGTATGTGCTGGCCAC
Q P F F D G S G T R R V D L P T Y A F Q R E R Y W L D A T Y T T G D

ATGCCGGCGGATCGGGCTCGGTTCCGGCGAGCATCCGTTGTTGGGTGCGGCTGTTGCTGTGGGTGGCAGGATGGTGTGCTGCTGACTGGTCTCTTTC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 5400
TACGCGCGCGCTAGCCCGAGCCAAGCCGCTCGTAGGCAACAACCCACGCCGACAACGACACCCACCGTGCCTACCACACGACGACTGACCAGCAGAAG
A A A I G L G S A E H P L L G A A V A V G G T D G V L L T G R L S

GGTTCAGACGCATCCGTGGCTGGCTGAGTATGTGGTTCAGGGTTCGGCGCTGCTGCCTGGTGCGGCTTCGTCGAGCTGGCGATCCGCGCGGGGATCAG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 5500
CCAAGTCTGCGTAGGCACCGACCGACTCATAACCAAGTCCCAAGCCGCGACGACGGACCACGCCGAAGCAGCTCGACCGTAGGCGCGCCCCCTAGTC
V Q T H P W L A E **Y V V Q G S A L L P** G A A F V E L A I R A G D Q
DH

GTCGGGTGCGATCTGGTGGAGGAGTTGACGCTTGAGGCGCCGCTTGCTCCTGCCTGTGCTGCGGTGCGGGTGCAGTTGTGGTGGTGGTGCCTCGGATG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 5600
CAGCCCACGCTAGACCACCTCCTCAACTGCGAACTCCGCGCGGAACAGGACGGACACAGACCCAGCCACGCCACGTC AACACCCAACCACGACGCCTAC
V G C D L V E E L T L E A P L V L P V S G S V R V Q L W V G A S D G

GGCGGGCAAGCGTGAAGTACGTTCTATCCGGTGGGGTATGTTGAGGACGGGCGTTCCTGGATCCGGCATGCCGTGGGTGTGCTGCGCGAGGGTGG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 5700
CCCGCCGTCGCACTCGACTGCAAGATAAGGCCACGCCACTACAACCTGCCCCGAAGGACCTAGGCCGTACGGCACCCACACGACGCGCTCCCACC
A G K R E L T F Y S G A G D V E D G R S W I R H A V G V L R E G G

CCGTTCTGAGGGGTTCCCTGGAGGTGTGGCCTCCGGCTGGGGCTGAGGTTGTGGATCTGGAGGGTTCACGACGAGGGGGCTGACGCCGGTCTCGCC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 5800
GGCAAGACTCCCCAAAGGACCTCCACACCGGAGGCCGACCCGACTCCAACACCTAGACTCCCAAGATGCTGCTCCCCGACTGCGGCCAGAGCGG
R S E G V S L E V W P P A G A E V V D L E G F Y D E G A D A G L A

TACGGCCGGTGTTCAGGGTTTGGCGCCGCGTGGCGCAAGGACGGCGAGGTGTTCCCGAGGTGCGGTTGCCGATGGTGTGAAGTCGGACGGCTTCG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 5900
ATGCCCGGCCACAAGGTCCCAAACGCCCGGCGCACCGGCTCCTGCGGCTCCACAAGCGGCTCCAGCGCAACGGGCTACCACACTTCAGCCTGCCGAAGC
Y G P V F Q G L R A A W R K D G E V F A E V A L P D G V K S D G F G

GTCTGCATCCGGCGCTGCTTGACGCGACGTTGCACACGGTCGGTCTCGCGACAGCGCGGACGTCGACGGCAAGGTGCCGTTCCGCTGGTGGTGGTGG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 6000
CAGACGTAGGCCGCGACGAAGTGCCTGCAACGTGTGCCAGCCAGAGCGCTGTCGCGCCTGCAGCTGCCGTTCCACGGCAAGCGCACCCAGCCACACGC
L H P A L L D A T L H T V G L A D S A D V D G K V P F A W S G V R

GCTGCACGCCTCCGCCGCGACGGCACTGCGGGTCCGGTTGACGCCCAAGAGGGGACGGCGTGACCTTGGCACTCGCCGACGGTTCGGGGCGCCCCGTTG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 6100
CGAGTGCAGGCGCGGCTGCCGTGACGCCAGGCCAAGTGCAGGTTGCTCCGCTGCCGACTGGAACCGTGAGCGGCTGCCAAGCCCGCGGGGGC
L H A S A A T A L R V R L T P T E A D G V T L A L A D G S G A P V

Anexo II

GCCACCATCGACTCCCTGGTGTGCTGCGCCCCGTCTCGCTCGACCAGATCGCCGGGGACGAGCGCGGCACGCGCTGTTCGGCGTTCGACTGGATATCCGCAC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 6200
CGGTGGTAGCTGAGGGACCACGACGCGGGGCAGAGCGAGCTGGTCTAGCGGCCCTGCTCGCGCCGTGCGCGACAAGCCGAGCTGACCTATAGCGTGT
A T I D S L V L R P V S L D Q I A G D E R G D A L F G V D W I S A P

CGCCACCGACGGCGCGCGCCGACGACCCAGTGGACCGACTGGACGCACTGGCGCAGTCGGCGGAAGTCCCGACTTCGTGCTCCTGCCGTGCCGAC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 6300
GCGGGTGGCTGCGCGCGCGCGGCTGCTGGGTACCTGGCTGGACCTGCGTGACCCGCTCAGCCGCTTACGGGCTGAAGCAGCAGGACGGCACGGGCTG
P T D G A A P T T Q W T D L D A L A Q S A E L P D F V V L P C P T

GACCTCCGCGTCCGACCTCGTTGCCGACGCGCACGAGACGGCGCACTGGCGCTGACCACGGTCCAGAAGTGGCTGGCCGACGACCGGTTCCGACGCCGCC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 6400
CTGGAGGCGCAGGCTGGAGCAACGGCTGCGCGTGTCTGCGCGTACCCGCGACTGGTGCAGGTCTTGACCGACCGGCTGCTGGCCAAGCTGCGGCGG
T S A S D L V A D A H E T A H W A L T T V Q N W L A D D R F D A A

CGCCTCGTCTTGTGACAGCGGCGGATCGCCACACGACCGACGAGGACGTACGGAACCTCGCGCAGGCCCGCTGTGGGGCCTGGTGGCGACGGCCC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 6500
GCGGAGCAGGAACACTGGTGCAGCGCTAGCGGTGGTGTGGCTGCTCCTGCATGCCTTGGAGCGCTCCGGCGGCACACCCCGACACCGGTGCCGGG
R L V L V T S G A I A T T T D E D V R N L A Q A A V W G L V R T A Q

AGTCGGAGAACCCCGACCGCATCGTGTGCTGCTGACCTCGACGACGAGGCGGCTCGCGCGACGCCCTTGCCGTCGCCCTGGCCACCGGCGAGCCCCAACT
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 6600
TCAGCCTCTTGGGGCTGGCGTAGCACGAGCAGCTGGAGCTGCTGCCCGGAGCGCGTGCAGAACGGCAGGCGGGACCGGTGGCCGCTCGGGGTGA
S E N P D R I V L V D L D D E A A S R D A L P S A L A T G E P Q L

CGCCGTACGCTCCGGCGACATCCGCACACCCCGCTGGCCAGGCTCGGCCGCTCGGCAGTGGACGACCCCGGCTTCGGCTCGGGCGGGTGTGTTGTCACC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 6700
GCGGCATGCGAGGCCGCTGTAGCGGTGGGGCGGACCGTCCCAGCCGGCAGCCGTACCTGCTGGGGCCGAAGCCGAGCCCGCCACGACAGTGG
A V R S G D I R T P R L A R V G P S A V D D P G F G S G A V L V T

GGTGCCACGGGCACGCTGGGCGGCCTCATGGCCCGCATCTCGTGCAGGCGGGTGTACGCAACCTGCTGCTGGTGGAGAGGCGGGCGCCACCGCCG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 6800
CCACGGTGCCTGCGACCCGCGGAGTACCGGGCCGTAGAGCAGCGGCTCGGCCACATGCGTTGGACGACGACCACTCGTCCGCGCCGGTGGCGGC
G A T G T L G G L M A R H L V A E R G V R N L L L V S R R G A T A E
KR

AAGGCGCCGAGGAAGTGCAGCGCCGAACTCGCCGCGAGGGCGCCGAAGTACGCTACGCCCTGCGACGTCGCCGACCGGGAAGCCGTCGCCGCACTGCT
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 6900
TTCCGCGGCTCCTTACGCGCGGCTTGGAGCGGCGGCTCCCGCGGCTTACTGCATGCGGCGGACGCTGCAGCGGCTGGCCCTTCGGCAGCGGCTGACGA
G A E E L R A E L A A Q G A E V T Y A A C D V A D R E A V A A L L

CGCCGAGCACAAGGTCTCCGCCGCTGTCACACCGCCGGTGTCTCGACGACGGCACCATCGGCTCACTGACAGCGGACCGGATCGACACCGTCTTCGG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 7000
GCGGCTCGTGTTCAGAGGCGGCGACGCTGTGGCGGCCACAGGAGCTGCTGCCGTGGTAGCCGAGTACTGTCGCTGGCCTAGCTGTGGCAGAAGGCC
A E H K V S A V V H T A G V L D D G T I G S L T A D R I D T V F R

CCCAAGGTCGACGCGGCTGGCACCTGCACGAGCTGACCCGCGACCTCGACCTGTCCGCGTTCGTCCTGTTCTCGTCCGTCGCGGTTGTTTTGGCGGGC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 7100
GGGTCCAGCTGCGCCGACCGTGGACGTGCTGACTGGGCGCTGGAGCTGGACAGGCGCAAGCAGGACAAGAGCAGGACGCGGCCACAAAAGCCCGCGG
P K V D A A W H L H E L T R D L D L S A F V L F S S V A G V F G G P

Anexo II

CCGGACAGGGCAACTACGCCGCCGGAACACCTTCCTGGACGCCCTCGCCCAGCACCGCCGCGCCACCGGCCTGCCCGCCACCTCACTCGCGTGGGGCCT
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 7200
GGCCTGTCCCCTTGATGCGCGCGCCTTGTGGAAGGACCTGCGGGAGCGGGTCTGGCGCGCGGTGGCCGACGGGCGGTGGAGTGAGCGCACCCCGGA
G Q G N Y A A A N T F L D A L A Q H R R A T G L P A T S L A W G L

GTGGGCGGACCGCGCGCATGGCGGGCAGCGTCGACGAGGCCGACGTGAAGCGGATGAGCCGGGGCGCGTGATCCCGCTCGGCTCCGCCAGGGCCTG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 7300
CACCCGCCTGCGGCCCGCTACCGCCCGTCGCAGCTGCTCCGGCTGCACTTCGCCTACTCGCCCGCCGCACTAGGGCGAGCCGAGGCGGGTCCCGGAC
W A D A G G M A G S V D E A D V K R M S R G G V I P L G S A Q G L

GCGCTGTTTCGACGCGGCCCGCACCCCGGGCGCGCCACGCTGGCCCCGTCGCGCTCGACCTGCCCGCCCTGCACCGGCAGGCCCGCACGAGCCCCGTAC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 7400
CGCGACAAGTGTGCGCCGCGCTGGTGGCCCCGCGCGGTGCGACCGGGGACGCGCGAGCTGGACGGGCGGACGTGGCCGTCCGGGCGTGTCTCGGGCATG
A L F D A A G T T G R A T L A P V P L D L P A L H R Q A R T S P V P

CGCACCTGCTGCGCGGTCTGGTGC GCGGCACGGCACGCAGGACCGCCGAGTCCGGCGCCGCCACCGGGTCCGCCCTCGCCCAGTCCCTTGGGGCCTGAC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 7500
GCGTGGACGACGCGCCAGACCACGCGCCGTGCCGTGCGTCTGGCGGCTCAGGCCGCGCGGTGGCCAGGGCGGAGCGGGTCAAGGAAACGCCCGGACTG
H L L R G L V R G T A R R T A E S G A A T G S A L A Q S L A G L T

CGAGCGGAGCGGGAGAAGCGCTCCTCGACCTCGTCTGGAGCAGCTCGCCGTGGTCTCGGCCACGGCTCGGCGGGCGCCATCGTGCCCCACCGCGCG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 7600
GCTCCGGCTCGCCCTCTTCCCGGAGGAGCTGGAGCAGGACCTCGTGCAGCGGCACCAGGAGCCGGTGCCGAGCCCGCCGCGGTAGCACGGGTGGCGCGC
E A E R E K A L L D L V L E H V A V V L G H G S A G A I V P D R A

TTCAAGGAGCTCGGCTTCGACTCGCTGACCGCGGTGCAACTGCGCAACCGCTGGGCGGGCCACCGAACTGCGCCTGTCCGCGACGCTGGTCTTCGACC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 7700
AAGTTCCTCGAGCCGAAGCTGAGCGACTGGCGCCAGCTTGACGCGTTGGCGGACCCGCGCCGGTGGCTTACGCGGACAGCGCTGCGACCAGAAGCTGG
F K E L G F D S L T A V E L R N R L G A A T E L R L S A T L V F D H
ACP

ACCCGAGCCCGCTGGCGCTCGCGACGCACCTCAGCACCGAACTCATCGGGCACCCCGCCGTCGTGACGACCACCCGGGCACGGGTGGCGTGGAGCGCA
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 7800
TGGGCTCGGGCGACCGGAGCGCTGCGTGGAGTCTGGCTTGAAGTAGCCGCTGGGCGGCAGCACTGCTGGTGGGCCGTGCCAGCCGCACTCGCTGCT
P S P L A L A T H L S T E L I G A P A V V T T T R A R V G V S D E

GCCGATCGCCATCGTCGCGATGAGCTGCCGTTCCCGGGCGGGTCAACTCGCCGAGGACCTGTGGCAGCTCGTCTCCGACGGTGTGACGCGCGTGGC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 7900
CGGCTAGCGGTAGCAGCGCTACTCGACGGCCAAGGGCCCGCCAGTTGAGCGGGCTCTGGACACCGTCGAGCAGAGGCTGCCACAGCTGCGGCAGCCG
P I A I V A M S C R F P G G V N S P E D L W Q L V S D G V D A V G

GAGTCCCGCCAAACCGCGGCTGGGACCTGGCCAACGTGTACAACCCCGACCCGGACCACGCGGGCACCTCGTACGCCCGTGAAGGGCCTTCCCTACG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 8000
CTCAAGGGGCGGTTGGCGCCGACCTGGACCGGTTGCACATGTTGGGGCTGGGCCTGGTGCGCCGTGGAGCATGCGGGCACTTCCCGGAAGGAGATGC
E F P A N R G W D L A N V Y N P D P D H A G T S Y A R E G A F L Y D

ACCGCGACCAGTTCGACCCGACGTTCTTCGGGATCTCGCCGCGGAGGCCCTCGCCATGGACCCGACGAGCGGCTGCTCTGGAGACGTCGTGGGAGGC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 8100
TGGGCTGGTCAAGCTGGGCTGCAAGAAGCCCTAGAGCGCGCTCCGGGAGCGGTACCTGGGCGTGTGCGCCGAGGACCTCTGCAGCACCCCTCCG
A D Q F D P T F F G I S P R E A L A M D P Q Q R L L L E T S W E A

Anexo II

GTTCGAGCGGGCCGGATCGACCCGAGCGGGTGC CGGGCGCGACATCGGCGTCTTACC GGCTCGAACGTGCAGGACTACACCCTGCTCCTGGCGGAG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 8200

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

AGCGACGAGACCGTCAAGGCTTCATCGGCACGGGCAACGCGCCAGCGTGTGTCCGGCCGGCTGCCTACGCGTTTCGGCCTCGAAGGCCCGCCGTGA
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 8300

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

CCGTCGACACCGGCTGCTCGTCTGCTGGTGGCGCTGCACATGGCCGTGCAGGCACTGCGGACCGGTTGACTGCTCGATGGCGCTCGCGGGCGCGTGA
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 8400

GCGAGCTGTGGCGCACGAGCAGCAGCACGCCACCGGACGTTACCGGCACGTCCGTTGACGCTGGCCACTGACGAGCTACCGGAGCGCCCGCCGACTG
V D T A C S S S L V A L H M A V Q A L R T G D C S M A L A G G V T
KS

GGTGATGTCGACCGCGGCTCGTTCATCGAGTTTCAGCCGCGCAGCGGGACTCTCCACCGACGGCCGCTGCAGGCCGTTCTCGTCCGACGCGGACGGCACCC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 8500

CCACTACAGTTCGGCCCGGAGCAAGTAGCTCAAGTCGGCGGTCGCGCCTGAGAGGTGGTGC CGGCGACGTCGGCAAGAGCAGGCTGCGCCTGCCGTGG
V M S T P G S F I E F S R Q R G L S T D G R C R P F S S D A D G T

GGCTGGGGCGAGGGCGTCGGCACGGTCGTCCTTGAGCGCTGTCCGACGCCAGCGCAACGGGCACCAGGTGCTCGCCGTGGTACGGGCGTCCGCGATCA
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 8600

CCGACCCCGTCCCAGCAGCGTCCAGCAACTCGCGGACAGGCTGCGGGTTCGGTTCGCCGTTGGTCCACGAGCGGCACCATGCCCGCAGGCGCTAGT
G W G E G V G T V V L E R L S D A Q R N G H Q V L A V V R A S A I N

ACCAGGACGGCGCTCCAACGGCCTCACCGCCCCAACGGCCCCGCGCAGCAGCGGCTCATCCGGCAGGCACTGGCCGGCCGGCCCTCGCGCCGACCGA
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 8700

TGGTCCTGCGCGCAGGTTGCCGAGTGGCGGGGGTTGCCGGGCGCGTCGTCGCGCAGTAGGCCGTCGCGTAGCCGCGCGCGCCGAGCGCGGCTGGCT
Q D G A S N G L T A P N G P A Q Q R V I R Q A L A A A G L A P T D

CGTGGACCGGCTCGAGGCGCACGGCACCGGCACGAAGCTCGGGCACCCGATCGAGGCGCAGGCGCTCCTCGCCACCTACGGTCAGGACCGGGCCGCCGAC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 8800

GCACCTGCGGCGAGTCCGCGTCCGCGTGGCCGTGCTTCGAGCCGCTGGGTAGCTCCGCGTCCGCGAGGAGCGGTGGATGCCAGTCCGCGCGGCGGCTG
V D A V E A H G T G T K L G D P I E A Q A L L A T Y G Q D R A A D

CGGCCGCTGTGGCTGGGCTCGGTGAAGTCCAACATCGGGCACACCAGGGCGCCCGGTGTGCGGGCATCATCAAGATGGTCATGGCGATCGCGCACG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 8900

GCCGGCGACACCGACCCGAGCCACTTCAGGTTGTAGCCCGTGTGGGTCCC CGGCGGCCACAGCGCCCGTAGTAGTTCTACCAGTACCGCTACGCCGTCG
R P L W L G S V K S N I G H T Q G A A G V A G I I K M V M A M R H G

GCGTACTGCCGAAGACGCTGCACGTCGCGAGCCACCCCGAGGTCGACTGGTCCACCGGTTGATGTCGGCTCCTCACGAGGCCCCTCTGGCCGGA
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 9000

CGCATGACGGCTTCTGCGACGTGCAGCGGCTCGGGTGGGGGCTCCAGCTGACCAGGTGGCCACTACAGCGAGGAGTGGCTCCGGGTGAGGACCGCCT
V L P K T L H V A E P T P E V D W S T G D V A L L T E A H S W P D

CACCGACCGGCGCGCGCTTCGGCGTCTCGTCTCGGGTTCAGCGGCACCAACGCGCACGCCATCATCGAGCAGGCGCCGGTCCCGAAGCCGCCGAG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 9100

GTGGTGGCCGGCGGCGCAAGCCGAGAGCAGCAAGCCCAAGTCGCGTGGTGGCGGTGGTAGTAGCTCGTCCGCGGCGAGTGGCTTCGGCGGCTC
T D R P R R F G V S S F G F S G T N A H A I I E Q A P V T E A A E

Anexo II

CCGACCGCCGAACCGGACCCGTCCTCGTGCCGCTGCTCGTCTCGGCCAAGGACGACGAAGCACTGCGGGCGGGAGCGGAGCGGCTGCGCGCCACCTCG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 9200
GGCTGGCGGCTTGGCTGGGAGGACGACGGCGACGAGCAGAGCCGGTTCCTGCTGCTTCTGACGCCGCCCTCGCTCGCCGACGCGGGTGGAGC
P T A E P D P S L V P L L V S A K D D E A L R A G A E R L R A H L V

TGGCACACCCGAGCTGAACCCGCGGACGTCGCCAGCACGCTCGCCGTGGCGGGTTCGACAGCCGGGCGGTCTGCTGGGCGAGGACCGGA
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 9300
ACCGTGTGGGGCTCGACTTGGGCGCGTGCAGCGGTCTGCGAGCGGCACCGCGCCAGCCGGAAGCTGTCGGCCCGCCAGCAGCACCCGCTCCTGGCGCT
A H P E L N P R D V A S T L A V A R S A F D S R A V V V G E D R D

CGGCCTGCTCGCCGGCTCGGGCGCTCGTCAAGGACCGGCGGTGGCCGAGGCGTCAAGTTCGGCGGGTGGCGATGGGTGTCGCGCGCCCTGCG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 9400
GCCGACGAGCGGCCGAGCCGCGGAGCAGCTTCTGGTGCAGCCACCGGCTCCGAGCTCAGCCGCGCCACCGCTACCCACAGCGCGCCGAGCG
G L L A G L G A L V E G P T A V A R G V E S A R V A M G V A R P A

GGCAAGGTGGCGTTCGTGTTCCCGGTTCAGGGTTCAGTGGCTGGGTATGGGTGTTGGAGTTGGCGGAGTCGTCGCCGTTTCCTGCTCGTTGGAGG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 9500
CCGTTCCACCGCAAGCACAAGGGGCCAGTCCCAAGAGTCACCGACCCATACCCACACCTCAACCGCTCAGCAGCGGCCACAAGGCACGGAGCAACCTCC
G K V A F V F P G Q G S Q W L G M G V E L A E S S P V F R A S L E A

CGTGTGCGGAGGCGCTGGCGCCGTATGTCGACTGGTCTGTTGGGATGTGCTGGCTGACGCGCCGAGCTGGAGCGGGTTCGATGTTGGTGCAGCCCGCT
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 9600
GCACACGCTCCGCGACCGCGGCATACAGCTGACCAGCAACCCCTACACGACCGACTGCGCCGGCTCGACCTCGCCAGCTACACCACGTCGGGCGCGA
C A E A L A P Y V D W S L W D V L A D A A E L E R V D V V Q P A L

GTTCCTCGTATGGTGTGCTTGGCGGCTGTGGCGCTCGTACGGTGTGAGCCGGATGCGGTGGTGGGGCACAGTCAGGGTGGATCGCTGCGCGCTAT
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 9700
CAAGAGGCACTACCACAGCGAACCGCGCAACCCGCGAGCATGCCACAGCTCGGCCACCCACCCCGTGTGAGTCCCACTCTAGCGACGCCGATA
F S V M V S L A A L W R S Y G V E P D A V V **G H S** Q G E I A A A Y
ATa

GTCGCGGGTTCGCTGTCGCTTGGCGGACGCTGCCAAGATCGTGGCGCAGCGGAGCAAGGCCATTCTCGCGCTGTCGGGTCAGGGCGGCATGGTGTGCTCG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 9800
CAGCGCCACGCGACGCAACGCCTGCGACGGTTCAGCACCCTGCGCTCGTTCGGTAAAGAGCGGACAGCCAGTCCCGCCGTACCACAGCGAGC
V A G A L S L A D A A K I V A Q R S K A I L A L S G Q G G M V S L A

CGGCACGGTTCGACGACGCTCGTGAAGTGGATCGCGCCGTGGGCGAGCGGGTCTCGGTTGCCGCGGTCACGGGCTGGTTCGGTGGTGGCGGGTGA
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 9900
GCCGCTGCCAGCTGTCGAGCACTCACCTAGCGCGGCACCCGCTCGCCAGAGCCAACGGCGCCAGTTGCCCGGACCAAGCCACCACCACCGCCACT
A T V D D A R E W I A P W G E R V S V A A V N G P G S V V V A G D

CGCGACGCGCTGGACGAGCTGGTGGAGGCGTGTGCCCGATGGCCGGCGGGCCCGCGGATACCCGTCGACTACCGCTCGCACTCGCCGATGTCGAG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 10000
GCGCTGCGGACCTGCTCGACCACCTCCGCACACGGCGGCTACCGCGCGCCGGCGGCTATGGGAGCTGATGCGCAGCGTGGCGGCTACAGCTC
A D A L D E L V E A C A A D G R R A R R I P V D Y A S H S P H V E

CGCATCGAGGATGAGTTGGCGAGTGTGCTGGCCGGGTGGAGCCCCAGGTGCGCCACTGTGCCGATGCTGTCGACCGTACCGGGGAGTGGCTGACCGGTA
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 10100
GCGTAGTCTACTCAACCGCTCACACGACCGGCCACCTCGGGTCCAGCGGTGACACGGCTACGACAGCTGGCAGTGGCCCTCACCAGTGGCCAT
R I E D E L A S V L A G V E P Q V A T V P M L S T V T G E W L T G T

Anexo II

CCGAGGTGGGTGCCGATTACTGGTATCAGAACCTGCGGCACACGGTGCCTTCGAGGAGGCCACGCGGGTCTCCTCGACCACGGCATCGACGCCTTCGT
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 10200
GGCTCCACCCACGGTAATGACCATAGTCTTGGACGCCGTGTGCCACGCAAAGCTCCTCCGGTGCGCCAGAGGAGCTGGTGGCGTAGCTGCGGAAGCA
E V G A D Y W Y Q N L R H T V R F E E A T R G L L D H G I D A F V

CGAATGCAGCCCCACCCCGCACTGCCCTTCGGGATACGCGAGACGCTCGAAGCGACCGCGGGCGAAGCCGTGCTGCTCGGCACCCTGCGCCGCGACGAC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 10300
GCTTACGTCGGGGTGGGGCGTGACGGGAAGCCCTATGCGCTCTGCGAGCTTCGCTGGCGCCGCTTCGGCAGCAGCAGCCGTGGGACGCGCGCTGCTG
E C S P H P A L P F G I R E T L E A T A A E A V V V G T L R R D D

PstI
GGGAACCTCGACCGGTTCTGTCAGTCCGTCGCGGAGGCGTACGTGGCGGGCTGCCCTTCGACTGGCGCCAACCTGGTACCCGAAGGACAGCAGCTGGACC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 10400
CCCTTGGAGCTGGCCAAGGACGTCAGGCAGCGGCTCCGCATGCACCGCCCGGACGGGAAGCTGACCGCGGTTGACCATGGGCTTCCTGTGTCGACCTGG
G N L D R F L Q S V A E A Y V A G L P F D W R Q L V P E G Q H V D L

TGCCACCTACGCCTTCGGCGCGGGCGGATCTGGCCGAACATGGCCGCGGCCCCACCGGCGACGCCACCGGGTCCGCCGTCGACGCGCAGTTCTGGGC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 10500
ACGGGTGGATGCGGAAGGCCGCGCCCGCTAGACCGGCTTGTACCGCGCGGGGGTGGCCGCTGCGGTGGCCAGCGCGCAGCTGCGCGTCAAGACCCG
P T Y A F R R G R I W P N M A A A P T G D A T G S A V D A Q F W A

SacI
GGCCGTCGCGGACGGCGACCTCGCGGAACTCGCGCGCCGAGCTCGAGTGTCCCTCGACGACCCGATCGGCACCGTGTGCCCGGGTGGCGTCTGTCGGC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 10600
CCGGCAGCGCCTGCGCTGGAGCCGCTTGGAGCGGGCTCGAGCTGCACAGGGAGCTGCTGGGCTAGCCGTGGCAGCAGCGGGCCGACCGCAGCACCGCC
A V A D G D L G E L A A E L D V S L D D P I G T V L P R L A S W R

CAGGCCAGTCGGAGAACCTCGAGATCGACGGCTGGCGCTACCGGCTCACCTGGCTGCCGGTCCCGGAGCCACCGCCACGCTCACGGGCACCTGGCTCG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 10700
GTCCGGGTCAGCCTCTTGGAGCTTAGCTGCCGACCGCATGGCGCAGTGGACCGCGCCAGGGGCTCGGGTGGCGGTGCGAGTGGCCGTGGACCGGAGC
Q A Q S E N S E I D G W R Y R V T W L P V P E P T A T L T G T W L V

TGGTCAGCCGCGACGGCGCAGGTGAGGAATCCGTCGCGCGCCCTGACCCGGCACGGCGCCCGCTCTGCCGGTGCAGCGCCGAGAACCCCTGGACCG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 10800
ACCAGTCGGCGCTGCCGCTCCACTCCTTAGGCAGCGGCGCGGGACTGGGCGGTGCCGCGGGCGCAGGACGGCCACGCGCGCCTCTTCGGGACCTGGC
V S R D G A G E E S V A A A L T R H G A R V L P V R A G E A L D R

GGACGCGCTGACGGCCAAGCTCACGGTTCGACGAGCCCGTCCGCGGGCTGCTCTCGTGTTCGCCCTCGACCCCGGACTGGCGTCGGGCGGGCTGTACCG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 10900
CCTGCGGACTGCCGTTGAGTGCAGCTGCTCGGGCAGCGCCCGCACGAGAGCGACAAGCGGGAGCTGGGGCTGACCGCAGCCCGCCGACATGCGC
D A L T A K L T V D E P V A G V L S L F A L D P G L A S G G V Y A

SacI
ACCTTCGCCCTCGTCCAGGCCCTCGGCGACACCGAGCTCACCGTCCCCTGTTGGCAGCCACCCGGGGCGGCTCACGGTTCGGCCGGTCCGACACACCG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 11000
TGGGAGCGGGAGCAGGTCCGGGAGCCGCTGTGGCTCGAGTGGCGAGGGCACACCGTGCAGTGGGCCCCGCCGAGTCCAGCCGGCCAGGCTGTGTGGG
T L A L V Q A L G D T E L T A P V W H A T R G G V T V G R S D T P V

TCAGCCCGAGCAGGCGATGATCTGGGGCTCGGCCGAGTCTCGGCAAGGAACAGCCCGAGCGCTGGGGCGGGTCTCGTTCGACTGCCGGACACGCTGGA
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 11100
AGTGGGGCTCGTCCGCTACTAGACCCCGAGCCGCTCAGCAGCCGTTCTTGTGGGCTCGCAGCCCGCCAGAGCAGCTGGACGGCCTGTGCGACCT
S P E Q A M I W G L G R V V G K E Q P E R W G G L V D L P D T L D

CGATCGCGCCCGCCCGCTCGCGGGTGTCTCTCCCGCACGACGAGGACGAGGACAGGTCGCCCTCCGCGCGGGCGGTGTGTTCCGCCGACGAGCTGGTG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 11200
GCTAGCGCGCGGGCGGAGCGCCACAGGAGAGGGCTGCGTCTGCTCCTGGTCCAGCGGGAGGCGCGCCCGCCACACAAGCGGGCGTCCGACCAC
D R A A A R L A G V L S R T Q D E D Q V A L R A G G V F A R R L V

Anexo II

CAGGCCCCCGTGGGGCACCTCGGGGACTGGGACCCACCGGCACGGTGTGGTACCGGGCGCACCGGGGGATCGGCGCGCACGTGGCCAGGTGGCTGG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 11300
GTCCGGGGGACCCCGTGGAGCCGCTGACCCTGGGGTGGCGGTGCCACGACCAGTGGCCGCCGTGGCCGCCCTAGCCGCGCGTGCACCGGTCCACCGACC
Q A P V G H L G D W D P T G T V L V T G **G T G G I G A H V A** R W L A
KR

CCGCGAGGGGCGCCGAGCACCTGCTGCTCACCAGCAGCGCGGGCCGACGCCGACGGCGGGAGGAAGTGCAGCGCGCAACTGACCGGGCTCGGCGCGCA
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 11400
GGCGTCCCCCGGGCTCGTGGACGACGAGTGGTCTCCGCGCCCGGCTGCGGCTGCCGCGCTCCTTGACGCGCGGCTTGACTGGCCCCGAGCCGCGCT
A R G A E H L L L T S R R G P D A D G A E E L R A E L T G L G A Q

GGTGACGATCGCCGCTGCGACGTGCGCCGACCGCGACGCGCTGGCGAAGCTGCTGGCGTCCGTTGCCGGAGGAGTCCCGCTCACCGCCGTATGCACAG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 11500
CCACTGCTAGCGGGGACGCTGCAGCGGCTGGCGTGCAGCGACCGCTTCGACGACCGCAGCCACGGCTCCTCAAGGGCGAGTGGCGGCAGTACGTGTGC
V T I A A C D V A D R D A L A K L L A S V P E E F P L T A V M H T

GCGGGCATCCTGGACGACGGCGTGTCTCACCAGCGCTCACCCCGGAGCGGCTCGCCGCGTGTGCGGCGGAAGGTCGAGGCCGCGATCCACCTGGACGAGC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 11600
CGCCCGTAGGACCTGCTGCCGACGAGTGGCGGAGTGGGGCCTCGCCGAGCGGGCGGCACGACCGCGGCTTCCAGCTCCGGCGCTAGGTGGACCTGCTCG
A G I L D D G V L T A L T P E R L A A V L R P K V E A A I H L D E L

TGACCAGGGACCAGGAGCTGTGGCGTTCGTGCTTCTCCTCGATCTCCGGCAGTTCGGCGCGATCGCCAGGGCAGTACTCGGCCGCCAACGCCTT
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 11700
ACTGTTCCCTGGTCTCGACAGCCGCAAGCAGACAAGAGGAGCTAGAGGCCGTGCAAGCCGCGCTAGCGGGTCCCCTCGATGAGCCGGCGGTTGCGGAA
T R D Q E L S A F V L F S S I S G T F G A I A Q G S Y S A A N A F

CCTCGACGCCTTCGCGCAACGGCGTGTGGCGCGGGCCTGCCCGCCACCTCGGTGGCGTGGGGCCCGTGGGGCGGGCGCGCATGGTCCGGCGAGAAGCTC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 11800
GGAGCTGCGGAAGCGGCTTGCCGCACACCGGCGCCCGGACGGGCGGTGGAGCCACCGCACCCCGGGCACCCCGCCGCGCGTACCAGCCGCTTTCGAG
L D A F A E R R V A A G L P A T S V A W G P W G G G G M V G E K L

GCCGAGATCCTGCGCAGCCGCGCACCCGCCCGTTCGAGCCCGAGCTGGCCATCGCCACGCTGGACCGCGCCCTGCGCCAGGACGGCCCCGCGCTCCTCG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 11900
CGGCTCTAGGACGCTCGGCGCCGTTGGCGGGGCAAGCTCGGGCTCGACCGGTAGCGGTGCGACCTGGCGCGGGACGCGGTTCTCGCGGGGCGGCAGGAGC
A E I L R S R G T A P F E P E L A I A T L D R A L R Q D G P A V L V

TCGTGGCGGACATCGACTGGAACGTGTACGCGCCGTCGCTGCCCGCCTCGCTGCTGCGGGACCTGCCCGCCGCGCAAGGCCCGGGCCCCCGTTCGAGGA
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 12000
AGCACCGCCTGTAGCTGACCTTGACATGCGCGGACGACGGGGGAGGACGACGACGCCCTGGACGGCGCGCGGTTCCGGGGCGGGGGGAGCTCCT
V A D I D W N V Y A P S L P A S L L R D L P A A R K A P A P V E E

GGCGGCGCCGAGGACCTCGCGGCCGAAGTCCGCGCACTGCCAGGGCCGACCGGGAGCGGCTCGTGTGGACCTGGTGCAGGACAGGTTCGCGCGGCT
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 12100
CCGCCGCGGCTCCTGGAGCGCCGCTTGAGCGCGGTGACGGTCCCGGCTGGCCCTCGCGCAGCAGACCTGGACCACGCGCTCGTCCAGCGGGCGGAG
A A P Q D L A A E L A A L P G P D R E R V V L D L V R E Q V A A V

CTCGGCTACGACAGCAGCACCCAGGTGGAACCCGCGGGCGTTCAGGACCTGGGCTTCGACTCGCTCACCAGGTCGAGTTCGCAACGTGTTCAACT
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 12200
GAGCCGATGCTGCTGCTGGTCCAGCTTGGGCCGCGCCGCAAGTTCCTGGACCCGAAGCTGAGCGAGTGGCGCCAGCTCAAGGCGTGCACAAGTTGA
L G Y D S S T Q V E P G R A F K D **L G F D S L T A V E** F R N V F N S
ACP

ACCTCGCCGTGCACGCGCTCCGGCACAGGGAGTGCGGGCTCGCCCTCGCGGGCGCGTACCCTGATGTGACGCCCCGACACCTTCGTGAGTTAGCCG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 13300
TGGAGCGGCACGTGCGCGAGGCCGTGTCCCTCACGCCGAGCGGGAGCGCCCGCAGTGGCACTACAGCTGCGGGCTGTGGAAGCAGCTCAAGTCGGC
L A V H A L R H R E C G L A L A G G V T V M S T P D T F V E F S R

CCAGCGGGCTCGCCGAGAACGGCCGCTGCAAGCCCTTCGCCGACGCGCGACGGCACCGGCTGGGGCGAGGGCGGGCCTCCTGCTCCTTGAGCGC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 13400
GGTCGCGCCGGAGCGGCTCTTGCCGGCGACGTTCCGGGAAGCGGCTGCGGGCGGCTGCCGTGGCCGACCCCGCTCCCGCGCCCGGAGGACGAGGAAGTCCGGC
Q R G L A E N G R C K P F A D A A D G T G W G E G A G L L L L E R

CTGTCCGACGCCAGCGCAACGGCCACCCGATCCTCGGGCTGGTACGCGGTTTCGGCGGTCAACCAGGACGGCGCGTCCAACGGCCTCACCGCCCCGAACG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 13500
GACAGGCTGCGGGTTCGGTTCGGGCTAGGAGCCGACCATGCGCCAAAGCCGACGTTGGTCTGCGCGCAGGTTGCCGAGTGGCGGGGCTGCG
L S D A Q R N G H P I L G V V R G S A V N Q D G A S N G L T A P N G

GCCCGTCGACGACGCGCTCATCCGCCAGGCCCTGGCCACTCGGGCCTCGCACCTCCGACGTGGACGCCGTTCGAGGCACACGGCCACGGGCACGACGCT
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 13600
CGGGCAGCGTCGTGCGCAGTAGGGGTCGGGACCGGGTAGCCCGGAGCGTGGGAGGCTGCACCTGCGGCAGCTCCGTGTGCCGTGCCCGTGTGCGA
P S Q Q R V I R Q A L A H S G L A P S D V D A V E A H G T G T T L

CGGTGACCCATCGAGGCCAGGGCTCCTCGCCACCTACGGCCAGAACAGGGCCGCGACCGGGCGCTGTGGCTGGGCGCCCTGAAGTCGAACATCGGC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 13700
GCCACTGGGTAGTCCGGGTCCGGGAGGAGCGGTGGATGCCGGTCTGTCCCGGGCTGGCCGGCGACACCGACCCGCGGGACTTCAGCTGTAGCCG
G D P I E A Q A L L A T Y G Q N R A A D R P L W L G A L K S N I G

CACACCCAGGCGGTTTCGGCGTCTGCGGTGATGAAGATGGTCTCGCGATGCAGCACCGGCTGCTGCCGAAGACCTCCACGTGGACAAGCCGCTCCT
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 13800
GTGTGGGTCCGCCAAAGGCCGACGACCGCACTACTTCTACCAGGAGCGCTACGTCGTGGCGCTCGAGCCGTTCTGGGAGGTGCACCTGTTCCGGCAGGA
H T Q A V S G V A G V M K M V L A M Q H R V L P K T L H V D K P S S

CGCACGTGACTGGTCGGCGGGCGCGGTGGAGCTGCTGACCGAGGCCCGGTCGTGGCCGGAGACCGGCAGGCCGTGGCGAGCGGGTGTGTCGTCGTTCCGG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 13900
GCGTGCAGCTGACCGCCCGCCGCCACCTCGACGACTGGCTCCGGGCCAGCACCGGCTCTGGCCGTCGGCACCGCTCGCCACACAGCAGCAAGCC
H V D W S A G A V E L L T E A R S W P E T G R P W R A G V S S F G

CGTCAGCGGCACCAACGCGCACGTCATCGTCAACAGGCTCCTGTCAATGAGCAG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 13955
GCAGTCGCCGTGGTTGCGCGTGCAGTAGCAGCTTGTCCGAGGACAGTTACTCGTC
V S G T N A H V I V E Q A P V N E Q

ANEXO III

Secuencia de nucleótidos de 2850 pb aisladas del cromosoma de *Streptomyces* sp. CO-AC25 que alberga el extremo 5' terminal del gen *orf7* así como los genes completos *orf8* y *orf9*. Se muestra la secuencia de aminoácidos deducida de cada región abierta de lectura. En negrita se marca la secuencia de nucleótidos codificante y con flechas los inicios de transcripción de cada gen.

Anexo III

CGTCTGGAGGAAGCCGCCTTCGCGGACGTAGCTGGTGCCGGGCCGGTTCGGGGTTCGGTGTCTAGAGGGCGGCGAGGTCCCAGCCGCGGTTCGGCGGGGA
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 100
CGCAGCACCTCCTTCGGCGGAAGCGCCTGCATCGACCACGGCCCGCCAGCCCCAGCCACAGCATCTCCCGCGCTCCAGGGTTCGGCGCCAGCCGCCCTT
A D H L F G G E R V Y S T G P R D P D T D Y L A A L D W G R D A P

AGCCGGAGATGGCGTCGACGCCGTTCGGCGACCAGTGCACAGCTCCTCGGGGGTTTCGACGCCCGCGGCGAGGCGGCGAGCTCATGGCGACGATGGCGAT
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 200
TCCGCCTCTACCGCAGCTGCGGCAGCCGCTGGTTCGACGGTTCGAGGAGCCCCAAAGCTGCGGCGGCCGTCGCCCGTTCGAGTACCGCTGCTACCGCTA
F G S I A D V G D A V L Q W L E E P T E V G G P L R C S M A V I A I

GGGCTCGTCTGGCGCGCCGGGCCGTGGCGGGTACGTAGCCGGCGTTCGGCGCTGTCAAGAGTTTCGGCGCGCAGATGGCGGGCGAGGCTGTTCGGTGTCT
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 300
CCCGAGCAGCAGCCGCGGGCCCGCACCCCATGCATCGGCCGACGCCGACAGCTTCTCAAGCCGCGGTCTACCGCCGCTCCGACAGCCACGAG
P E D D A R R A T A P V Y G A D A S D F L E A R L H R A L S D T S

GGGTAGTCAAGACGGCGGTGACGGGCAGCTTGAAGCCGGTGGCGCCCTGCAGACGGCGGTGCAGCTCCACCGCGGTTCAGGGAGTCCAGGCCAGGTTCGA
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 400
CCCATCAGCTTCTGCCGCCACTGCCCGTGAACCTCCGGCCACCGGGGACGCTCTGCCGCCACGTCGAGGTGGCGCCAGTCCCTCAGGTCCGGGTCCAGCT
P Y D F V A T V P L K L G T A G Q L R R H L **E V A T L S D L G L D**

ACP

GGAAGGGCGGGAGAGGTTCGCCGGGCGGGCGGGTGCCTCCGCGCCCGGAGGACCGCGGACCTGCTCGGCCACCACGGCGCGCAGCCGCCGTTCTCTG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 500
CCTTCCCGCCCTCTCCAGCGGCCCGCCGCCACGGCAGGCGCGGGCTCCTGGCGCCGTCGACGAGCCGGTGTGTGCCGCGCTCGGGCGGC **AAGGA**
L F P R S L D G P P R T G D A G G L V A A V Q E A V V **RBS**

orf7 ←

EcoRI

TTCCGCGCGGGGCGGCCTTCAAGGGTGTGCGCAGCTCTTCCGGGTGAGCCTCCGGGTGAGGCCGGGCTCCGGGTGAGAAATGGGCTGC**GAATTC**CTCG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 600
AAGGCGCGCCCGCCGGAAGTTCCACAGCGCGTTCGAGAAGGCCACTCGGAGGCCACTCCGGCCCGGAGGCCACTTACCAGCCTTAAGGAGC

AAGTGCCTCATTGTTCCACTCCGTAGCACCGGCCATGATGCTTTCAACAAATCCCTGACTGCACTCCGGGCGAAATCATCTGACGGCGCGGGGACGC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 700
TTCAGCGAGTAACAAGGTGAGGCATCGTGGCCGTTACTACCGAAAGTTGTTTAGGGACTGACGTGAGGCCCGCTTTAGTAGACTGCCGCGCCGCCCTGCG

ACGGAACCAGGAATCCGGGACTCGTCGGTGATTCCCCGAGAAGCTGCGAACTCCTGTGACCGCACGGATTCTCTGAGCGGATTACTGAAATGTTTCAGGA
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 800
TGCCTTTGGTCTTAGGCCCTGAGCAGCCACTAAGGGGCTCTTCGACGCTTGGAGACACTGGCGTGCCTAAGGACTCGCCTAATGACTTTACAAGTCTT

ACACCGACATACCGACCGGCTTCAGACAAATTTGCCACAATACCTGTGACGGCTAACAGTGCAGCCTTCGTCGCCCCTACCCCTAAGCGGCCCTAT
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 900
TGTGGCTGTATGGGCTGGCCGAAGTCTGTTTAAACGGGTGTTATGGACACGTCCGATTGTCACGCCGAAGCAGCGGGGATGGGGATTCCCGGGGATA

RBS

GGCGCCCGATGCGGGCGGATATAGGGTGTTCGGGATTCTCGCCCGCCCTAGGGTTCGGGATCGTGAAAACGCTGTCCGGGTT**GAGAA**TCTGT**GTC****AA**
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 1000
CCGCGGGCCTACGCCCGCCTATATCCCCACAAGCCCTAAGAGCGGGCGGGATCCAGCCCTAGCACCTTTTGCACAGGCCCAACTCTTAGACCACTT

V K
└─┬─> **orf8**

ACAGTACCGCTCAACGCACGCCGCTGGTTCGACGGCCTGAGCTTCGATGTGTATCGGGGCGAGGTGTTTCGACTGCTCGGCCCAATGGCGCGGGGAAA
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 1100
TGTATGCGCAGGTTGCGTTCGGCGCACCACTGCGGACTCGAAGCTACACATAGCCCCGCTCCACAAGCCTGACGAGCCGGGGTTACCGCGCCCTTT

Q Y A S N A R R V V D G L S F D V Y R **G E V F G L L G P N G A G K**
Walker A

AGCACCACGGTGGGGATGATGACCACCCGGGTGCGGCCGACCTCGGGCAGCGTCGTGATCGACGGTGTGACGCTGGTTCGCGCGACCCGAGAGAGGGCGAAAC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 1200
TCGTGGTGCACCCCTACTACTGGTGGGCCACGCGGGCTGGAGCCCGTTCGACGACTAGCTGCCACAGCTGCACCAGCGCGCTGGCTCTCTCCGCTTTG
S T T V G M M T T R V R P T S **G** S V V I D G V D V V A R P R E A K R

CGGTCCTGGCGGTGGTGCAGCGCAACAACCTCGACCGGTGCTGACGGTGCCTCAGAACCTGTTGTTCCACGCCAAGTACCACGGGATCCCGGGTGC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 1300
CGCAGGACCGCCACCACGGCGTTCGGTGTGGAGCTGGCCAGCAACTGCCACGCGTCTTGGACAACAAGGTGCGGTTTCATGGTGCCTTAGGGCCACG
V L A V V P Q R N N L D R S L T V R Q N L L F H A K Y H G I P G A

CGAGCGGGCGCAGCGGGCCGATGAGCTCCTCGATCAGATGGGCCCTCGCGGACTACCSCAAGTCAAGGTGACGACATGTCCGGCGGGCAGGTGCAGCGC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 1400
GCTCGCCCGCTCGCCCGGCTACTCGAGGAGCTAGTCTACCCGGAGCGCCTGATGGCGTTCAGCTTCCAGCTGTGTACAGGCCCGCCGTCACGTCGCG
E R R Q R A D E L L D Q M G L A D Y R K S K V D D M **S G G** Q V Q R
lazo 3

GTGATGATCGCCCGCCCTGATCCACCGCCGAACGTGCTGTTCTCGACGAGCCCGCAGCGGGTCTCGACCCGCGAGCGCGCCTGTTCTGTCACGAGC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 1500
CACTACTAGCGGGCGCGGGACTAGGTGGCCGGCTTGACGACAAGGAGCTGCTCGGGCGCTGCCAGAGCTGGGCGTCCGCGCGGACAAGCAGTGTCTCG
V M I A R A L I H R **P N V L F L D E P A T G L D** P Q A R L F V H E R
Walker B

GGGTCGCCGAACCTCGCCGAGCGGGCGTGACGGTGGTCTCACCCACGACATGGACGAGGCCCAAGCTGTGCGACCGCTCGGCATCGTGGACCG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 1600
CCCAGCGGCTTGAAGCGCTCGCCCGCACTGCCACCAGCAGTGGTGGTGTGTACCTGCTCCGGCGGTTGACACGCTGGCGCAGCCGTAGCACCTGGC
V A E L A E R G V T V V V T T H D M D E A A K L C D R V G I V D R

GGGCACCCTGCTCGCCCTGGACACCCCGGAGGGCTCGGGCAAGATCCTGCCCGCGCACTCGACGCTGGCCCTGACGGTGGAGTCCCGCAGGGCCCGCG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 1700
CCCGTGGGACGAGCGGGACCTGTGGGGCTCCCGGAGCCGTTCTAGGACGGCGCGTGTAGCTGCGACCGGGACTGCCACCTCAGCGCGTCCCGGGCCCG
G T L L A L D T P E G L G K I L P A H S T L A L T V E S P Q G P A

CCGGTGGCGAAGCTGCTGTCCGACCTTCCCGGGTGGCGGAGGTCGACGAAGTGGACCCCGGAGAACCCCGCCGGAGGGCGGCTTCCGGTTCGGTATCT
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 1800
GGCCACCGCTTCGACGACAGGCTGGAAGGGCCCCACGCCGTCCAGCTGCTTACCTGGGCCTTCTGGGCGGGCGCCTCCCGCCGAAGGCCAAGGCATAGA
P V A K L L S D L P G V R Q V D E V D P E N P P A E G G F R F R I Y

ATACGGACACCGAGCCGGTGGCGGCGCTCCAGCCGCGCTCGCCGCGCTGACCGCGGGGACTTCCGGGTGACGGACGTCGCGCTCGGAAAGGTGTCCTT
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 1900
TATGCTGTGGCTCGGCCACCGCCGCGAGGTCGGCCGCGAGCGGCGGACTGGCCGCCCTGAAGGCCACTGCCTGCAGCCGCGAGCCTTCCACAGGGA
T D T E P V A A L Q P A L A A L T G G D F R V T D V G V G K V S L

GGAAGACGTCTTCTCCACCTCACCGAAAGAATCTGCG**ATGAC**TACGTGCGCCACGGCCGACCGGCTCGCGGAGTTGCACACGTTTCGGCGCCATCA
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 2000
CCTTCTGCAGAAAGAGGTGGAGTGGCCTTTCTTAGACGCTACTGATGCAGCCGGTCCCGGCTGGCCGGAGCGCCGTCACAGTGTGCAAGCCGCGGTAGT
E D V F L H L T G K N L R *
M T T S A T A A P A S R Q L H T F G A I M
└─┬─> ORF9

TGTGGCGGACGCTGTTGCTGCGCCCTGCGCCAGTTCACCGGTCATCGGCCAGGCCATCATCCAGCCGTTCTTACGCTGTTGCTGTTTCGGTGTGCTGCT
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 2100
ACACCGCGCTGCACAAGCAGCGGGACGCGGTCAAGGGGTGCAAGTAGCCGGTCCGGTAGTAGGTGCGCAAGAAGTGCACAAGCACAAGCCACACGACGA
W R D V F V A L R Q F P T F I G Q A I I Q P F F T L F V F G V L L

GAGCGACGCGGGCTATGTGGACGGCGACTACAGTGCCGTGCTGCTGCCCGGTGCTCTGGCCCTGAACGGATTCTCTCGGCGGCATGCAGAACACGGCACTG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 2200
CTCGCTGCGCCCGATACACCTGCCGCTGATGTACGGCACGACGACGGGCCACAGGACCGGGACTTGCCCTAAGGAGCCGCGTACGTCTTGTGCCGTGAC
S D A G Y V D G D Y S A V L L P G V L A L N G F L G G M Q N T A L

CCGCTCGTCTGGACTTCTCGTACACGCGGGAGATCGAGGACCGGCTCCTTTCGCCGATCTCGCTGAGCCTGGTGGCCGTCGAGAAGATTCTCTTCGGCG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 2300
GGCGAGCAGCACCTGAAGAGCATGTGCGCCCTCTAGCTCCTGGCCGAGGAAAGCGGCTAGAGCGACTCGGACCACCGGCAGCTCTTCTAAGAGAAGCCGC
P L V V D F S Y T R E I E D R L L S P I S L S L V A V E K I L F G A

CGCTGTACGGCTTCTGGCGTCTGATGATGACGCCGGTCCGATGATCCTTCTTGACCTGCTGCTGGCCGATCAGCAAGACCATTCCGGCTCTGGGCAT
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 2400
GCGACATGCCGAAGCACCGCAGCGACTACTACTGCGGCCAGCCGTAAGGAACTGGACAGCACCGGCTAGTCGTTCTGGTAAGGCCGAGACCCGTA
L Y G F V A S L M M T P V G M I L L D L S W P I S K T I P A L G I

CGCCGGGCTCGGCGCGCTCGCGGGTGCCTCGATCGGTCTGTGCCTCGGCACGTCCGGTGGAGCAGCGACAAGGTCGACGTGATGTTACGCTGGTCCCTGATG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 2500
GCGGCCGAGCCGCGGAGCGCCACGGAGCTAGCCAGACACGGAGCCGTCAGCCACTCGTCGCTGTTCCAGCTGCACTACAAGTGCAGCAGGACTAC
A G L G A L A G A S I G L C L G T S V S S D K V D V M F T L V L M

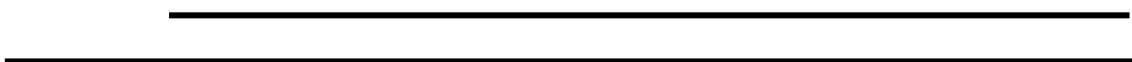
CCGCTGCTGTTACGGGCTCGACGAGTTCCTCGCTGGTCCGCTCGACGAGGTGCGCTGGTTCAGGTGATCTGCGGTGCCAATCCGCTGTCGTACGCCA
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 2600
GGCGACGACAAGTGCCCGAGCTGCGTCAAGGGCGACCAGCGGAGCTGCTCCACGGACCAAGGTCCACTAGACGCCACGGTTAGGCGACAGCATGCGGT
P L L F T G S T Q F P L V A L D E V R W F Q V I C G A N P L S Y A S

GTGAGGGCATGCGGGCCGTGACGTCCCGAACGTGGAGCACGTACCCTGCGGATCAGCTGATCGTCTGCTCGGGTGCATCAGCTGTTCCGGTCCAT
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 2700
CACTCCCGTACGCCCCGACTGCAGGGGCTGACCTCGTGCATGGCGACGGCTAGTGCAGTACGACGAGCCAGCTAGTGCAGACAAGCGCAGGTA
E G M R A V T S P N V E H V P L P I T L I V L V G S I T L F A S I

CGGTATCCGCGGGTTCCTCAAGCGCGCGCAGACGTGAGCGGAGGCCATCGGCCCTGCTCGCACGAAGGGCCCCGAAGCGGAGTCAGCCGCTTCGGGGC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 2800
GCCATAGGCGCCCAAGGAGTTCGCGCGCTGCTGCACTCGCCTCCGGGTAGCCGGGACGAGCGTGTCTCCCGGGGCTTCGCTCAGTCGGCGAAGCCCCG
G I R G F L K R A Q T *

CCTTCGTCCGTGCCCGCCGGTCACTCCTCGGGCAGGTCGACCGGGCC
-----+-----+-----+-----+-----+-----+ 2850
GGAAGCAGGCACGGGCCGGCCAGTCAGGAGCCCGTCCAGCTGGCCCCG

BIBLIOGRAFIA



Altschul, S.F., T.L. Madden, et al. (1997). "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs." Nucleic. Acids. Res. **25(17)**: 3389-3402.

Aparicio, J.F., I. Molnar, et al. (1996). "Organization of the biosynthetic gene cluster for rapamycin in *Streptomyces hygroscopicus*: análisis of the enzymatic domains in the modular polyketide synthase." Gene **169**: 9-16.

Aparicio, J.F., A.J. Colina, et al. (1999). "The biosynthetic gene cluster for the 26-membered ring polyene macrolide pimaricin: a new polyketide synthase organization encoded by two subclusters separated by functionalization genes." J. Biol. Chem. **274(15)**: 10133-10139.

Aparicio, J.F., R. Fouces, et al. (2000). "A complex multienzyme system encoded by five polyketide synthase genes involved in the biosynthesis of the 26-membered polyene macrolide pimaricin in *Streptomyces natalensis*." Chem. Biol. **7**: 895-905.

Bartolomé, B., Y. Jubete, et al. (1991). "Construction and properties of a family of pACYC184-derived cloning vectors, compatible with pBR322 and its derivatives". Gene **102**: 75-78.

Bentley, S.D., K.F. Chater, et al. (2002). "Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2)." Nature **417**: 141-147.

Berdy, J. (1995). "Are actinomycetes exhausted as a source of secondary metabolites?". *Proceedings of the 9th International Symposium on the Biology of Actinomycetes*. Part 1. Allerton Press, New York: 3-23.

Bevitt, D.J., J. Cortés, et al. (1992). "6-Deoxyerythronolide-B synthase-2 from *Saccharopolyspora erythraea*. Cloning of the structural gene, sequence analysis and inferred domain structure of the multifunctional enzyme." Eur. J. Biochem. **204**: 39-49.

Bibb M.J., J. White, et al., (1994). "The mRNA for the 23S rRNA methylase encoded by the *ermE* gene of *Saccharopolyspora erythraea* is translated in the absence of a conventional ribosome-binding site." Mol. Microbiol. **14**: 533-545.

Bierman, M., R. Logan, et al. (1992). "Plasmid cloning vectors for the conjugal transfer of DNA from *Escherichia coli* to *Streptomyces* spp." Gene **116(1)**: 43-49.

Birnboim, H.C. (1983). "A rapid alkaline extraction method for the isolation of plasmid DNA." Methods Enzimol **100**: 243-255.

Bisang, C., P.F. Long, et al. (1999). "A chain initiation factor common to both modular and aromatic polyketide synthases." Nature Lett. **401**: 502-504.

Bognár, R., B.O. Brown, et al. (1970). "The structure of Flavofungin". Tetrahedron Letters **Nº7**: 471-474.

Bolard, J. (1986). "How do the polyene macrolide antibiotics affect the cellular membrane properties?" Biochim Biophys Acta **864(3-4)**: 257-304.

Brautaset, T., O.N. Sekurova, et al. (2000). "Biosynthesis of the polyene antifungal antibiotic nystatin in *Streptomyces noursei* ATCC 11455: analysis of the gene cluster and deduction of the biosynthetic pathway." Chem. Biol. **7(6)**: 395-403.

Brautaset, T., S.E.F. Borgos, et al. (2003). "Site-specific mutagenesis and domain substitutions in the loading module of the nystatin polyketide synthase, and their effects on nystatin biosynthesis in *Streptomyces noursei*." J. Biol. Chem. **278**: 14913-14919.

Brown, M.G.M., A. Weston, et al. (1979). "Transformation of *E. coli* C600 by plasmid DNA at different phases of growth." FEMS Microbiol Lett **5**: 219-222.

Brown, D.P., S.J.D. Chiang, et al. (1988). "Site-specific integration in *Saccharopolyspora erythraea* and multisite integration in *Streptomyces lividans* of actinomycete plasmid pSE101." J. Bacteriol. **170**: 2287-2295.

Brown, M.J.B., J. Cortés, et al. (1995). "A mutant generated by expression of an engineered DEBS1 protein from the erythromycin-producing polyketide synthase (PKS) in *Streptomyces coelicolor* produces the triketide as a lactone, but the major product is the nor-analogue derived from acetate as starter unit." J. Chem. Soc., Chem. Commun. 1517-1518.

Bruton, C.J., E.P. Guthrie, et al. (1991). "Phage vectors that allow monitoring of secondary metabolism genes in *Streptomyces*". Bio/Technology **9**: 652-656.

Burke, R.C., J. Swartz, et al. (1954). "Mycoticin, a new antifungal antibiotic." J. Invest. Dermatol. **123**: 163-168.

Caffrey, P., S. Lynch, et al. (2001). "Amphotericin biosynthesis in *Streptomyces nodosus*: deductions from analysis of polyketide synthase and late genes." Chem. Biol. **8**: 713-723.

Campelo, A.B. and J.A. Gil (2002). "The candicidin gene cluster from *Streptomyces griseus* IMRU 3570." Microbiology **148**: 51-59.

Champness, W.C. and K.F. Chater (1994). "Regulation and integration of antibiotic production and morphological differentiation in *Streptomyces* spp. Regulation of Bacterial Differentiation. P. Piggot, C. P. J. Moran and P. Youngman. Washington, D. C., American Society for Microbiology: 61-93.

Chater, K.F. (1984). "Morphological and physiological differentiation in *Streptomyces*." *Microbial Development* (Shapiro, L. and Losick, R., ed), pp. 89-115. Cold Spring Harbor Laboratory (New York).

Combes, P., R. Till, et al. (2002). "The *Streptomyces* genome contains multiple pseudo-attB sites for the (phi)C31-encoded site-specific recombination system." *J. Bacteriol.* **184(20)**: 5746-5752.

Cortés, J., S.F. Haydock, et al. (1990). "An unusually large multifunctional polypeptide in erythromycin-producing polyketide synthase in *Saccharopolyspora erythraea*." *Nature* **348(6297)**: 176-178.

Cortés, J., K.E.H. Wiesmann, et al. (1995). "Repositioning of a domain in a modular polyketide synthase to promote specific chain cleavage." *Science* **268**: 1487-1489.

Covarrubias, L. and F. Bolivar (1982). "Construction and characterization of new cloning vehicles. VI. Plasmid pBR329, a new derivative of pBR328 lacking the 482-base-pair inverted duplication." *Gene* **17**: 79-89.

Cundliffe, E. (1978). "Mechanism of resistance to thiostrepton in the producing-organism *Streptomyces azureus*." *Nature* **272**: 792-795.

Cundliffe, E. (1989). "How antibiotic-producing organisms avoid suicide." *Annu Rev Microbiol* **43**: 207-233.

Deveraux, J., P. Haerberli, et al. (1984). "A comprehensive set of sequence analysis programs for the VAX." *Nucleic. Acids. Res.* **12**: 387-395.

Donadio, S. and C.R. Hutchinson (1991). "Cloning and characterization of the *Saccharopolyspora erythraea* *fdxA* encoding ferredoxin." *Gene* **100**: 231-235.

Donadio, S. and L. Katz (1992). "Organization of the enzymatic domains in the multifunctional polyketide synthase involved in erythromycin formation in *Saccharopolyspora erythraea*." *Gene* **111(1)**: 51-60.

- Donadio, S., J.B. McAlpine, et al.** (1993). "An erythromycin analog produced by reprogramming of polyketide synthesis." Prox. Natl. Acad. Sci. U.S.A. **90**: 7119-7123.
- Fernández-Moreno, M.A., E. Martínez, et al.** (1992). "Nucleotide sequence and deduced function of a set of co-transcribed genes of *Streptomyces coelicolor* A3(2) including the polyketide synthase for the antibiotic actinorhodin." J. Biol. Chem. **267**: 19278-19290.
- Fernández-Moreno, M.A., L. Carbó, et al.** (1998). "A silent ABC transporter isolated from *Streptomyces rochei* F20 induces multidrug resistance." J. Bacteriol. **180(16)**: 4017-4023.
- Frischauf, A.M., H. Lehrach, et al.** (1983). "Lambda replacement vectors carrying polylinker sequences." J Mol Biol **170**: 827-842.
- Gil, J.A., G. Naharro, et al.** (1985). "Characterization and regulation of p-aminobenzoic acid synthase from *Streptomyces griseus*." J. Gen. Microbiol. **131(6)**: 1279-1287.
- Gil, J.A. and J.F. Martín** (1997). "Polyene Antibiotics." Biotechnology of Antibiotics. New York, Marcel Dekker. 551-575.
- Guilfoile, P.G. and C.R. Hutchinson** (1991). "A bacterial analog of the *mdr* gene of mammalian tumor cells is present in *Streptomyces peuceticus*, the producer of daunorubicin and doxorubicin." Proc. Natl. Acad. Sci. USA **88**: 8553-8557.
- Hamilton-Miller, J.M.T.** (1973). "Chemistry and biology of the polyene macrolide antibiotics." Bacteriol. Rev. **37**: 166-196.
- Haydock, S., J.F. Aparicio, et al.** (1995). "Divergent sequence motifs correlated with the substrate specificity of (methyl)malonyl-CoA:acyl carrier protein transacylase domains in modular polyketide synthases." FEBS Lett. **374**: 246-248.
- Hazen, E. L. and R. Brown** (1950). "Two antifungal agents produced by a soil actinomycete." Science **112**: 423.
- Henderson, G., P. Krygsman, et al.** (1987). "Characterization and structure of genes for proteases A and B from *Streptomyces griseus*." J. Bacteriol. **169(8)**: 3778-3784.
- Hyde, S.C., P. Emsley, et al.** (1990). "Structural model of ATP-binding proteins associated with cystic fibrosis, multidrug resistance and bacterial transport." Nature **346**: 362-365.
- Hobbs, G., C.M. Frazer, et al.** (1989). "Dispersed growth of *Streptomyces* in liquid culture." Appl. Microbiol. Biotechnol. **31**: 272-277.

Hopwood, D.A., T. Kieser, et al. (1983). "Plasmids, recombination and chromosome mapping in *Streptomyces lividans* 66." J Gen Microbiol **129**: 2257-2269.

Hopwood, D.A., M.J. Bibb, et al. (1985). "Genetic manipulation of *Streptomyces*, a laboratory manual." The John Innes Foundation, Norwich, UK.

Hopwood, D.A., T. Kieser, et al. (1986). "The Bacteria." Vol. 9, pp. 159. Eds. Queener SW, Day LE. Academic Press (New York).

Hopwood, D.A., and D.H. Sherman (1990). "Molecular genetics of polyketides and its comparison to fatty acid biosynthesis." Annu. Rev. Genet. **24**: 37-66.

Hopwood, D. A. (1997). "Genetic contributions to understanding polyketide synthases." Chem. Rev. **97**: 2465-2497.

Hutchinson, C.R. and I. Fujii (1995). " Polyketide synthase gene manipulation: a structure-function approach in engineering novel antibiotics." Annu. Rev. Microbiol. **49**: 201-238.

Ikeda, H., J. Ishikawa, et al. (2003). "Complete genome sequence and comparative analysis of the industrial microorganism *Streptomyces avermitilis*." Nat. Biotechnol. **21(5)**: 526-531.

Itoh, Z., T. Suzuki, et al. (1984). "Gastrointestinal motor-stimulating activity of macrolide antibiotics and analysis of their effects on the canine gut." Animicrob. Agents. Chemother. **26**: 863-869.

Izumikawa, M., P.R. Shipley, et al. (2003). "Expression and characterization of the type III polyketide synthase 1,3,6,8-tetrahydroxynaphthalene synthase from *Streptomyces coelicolor* A3(2)." J. Ind. Microbiol. Biotechnol. **30(8)**: 510-515.

Janssen, G.R. and M.J. Bibb (1993). "Derivatives of pUC18 that have *Bgl*II sites flanking a modified multiple cloning site and that retain the ability to identify recombinant clones by visual screening of *Escherichia coli* colonies." Gene **124(1)**: 133-134.

Jiménez, A., M. Zalacaín, et al. (1987). "Clonación y análisis de genes implicados en la biosíntesis de antibióticos." *Nuevas tendencias. Ingeniería genética* (Renard, J. y Vicente, M., ed.): 496-502.

Kao, C.M., G.L. Luo, et al. (1995). "Manipulation of a macrolide ring size by directed mutagenesis of a modular polyketide synthase." J. Am. Chem. Soc. **117**: 9105-9106.

- Katz, L. and S. Donadio** (1993). "Polyketide synthesis: prospects for hybrid antibiotics." Annu. Rev. Microbiol. **47**: 875-912.
- Kessler, H. A., J. Dixon, et al.** (1981). "Effects of amphotericin B on hepatitis B virus." Antimicrob. Agents Chemother **20**: 826-833.
- Kieser, H.M., T. Kieser, et al.** (1992). "A combined genetic and physical map of the *Streptomyces coelicolor* A3(2) chromosome." J. Bacteriol. **174(17)**: 5496-5507.
- Kinashi, H. and T. Shimaji-Murayama** (1987). "Detection of giant linear plasmids in antibiotic producing strains of *Streptomyces* by the OFAGE technique." J. Antibiot. **40(6)**: 913-916.
- Kino, T., H. Hatanaka, et al.** (1987). "FK-506, a novel immunosuppressant isolated from a *Streptomyces*. II. Immunosuppressive effect of FK-506 *in vitro*." J. Antibiot. **40(9)**: 1256-1265.
- Kobinata, K., H. Koshino, et al.** (1993). "RK-397, a new oxo pentaene antibiotic." J. Antibiot. **46(10)**: 1616-1618.
- Kotowska, M., K. Pawlik, et al.** (2002). "Type II thioesterase from *Streptomyces coelicolor* A3(2)." Microbiology **148**: 1777-1783.
- Larson, J.L. and C.L. Hershberger** (1986). "The minimal replicon of a streptomycete plasmid produces an ultrahigh level of plasmid DNA". Plasmid **15**: 199-209.
- Leblond P., F.X. Francou, et al.** (1990). "Pulsed-field electrophoresis analysis of the genome of *Streptomyces ambofaciens* strains." FEMS Microbiol. Lett. **72**: 79-88.
- Leblond, P., M. Redenbach, et al.** (1993). "Physical map of the *Streptomyces lividans* 66 genome and comparison with that of the related strain *Streptomyces coelicolor* A3(2)." J. Bacteriol. **175(11)**: 3422-3429.
- Lezhava A., T. Mizukami, et al.** (1995). "Physical map of the linear chromosome of *Streptomyces griseus*." J. Bacteriol. **177**: 6492-6498.
- Lin, Y. S., H. M. Kieser, et al.** (1993). "The chromosomal DNA of *Streptomyces lividans* 66 is linear." Mol. Microbiol. **10(5)**: 923-933.
- Lomovskaya, N., K.F. Chater, et al.** (1980). "Genetics and molecular biology of *Streptomyces* bacteriophages." Microbiol. Rev. **44**: 206-226.

- Long, C.M., M.J. Virolle, et al.** (1987). "Alpha-amylase gene of *Streptomyces limosus*: nucleotide sequence, expression motifs, and amino acid sequence homology to mammalian and invertebrate alpha-amylases." J. Bacteriol. **169(12)**: 5745-5754.
- Lydiate, D.J., F. Malpartida, et al.** (1985). "The *Streptomyces* plasmid SCP2*: its functional analysis and development into useful cloning vectors." Gene **35(3)**: 223-235.
- Macneil, D.J., K.M. Gewain, et al.** 1992. "Analysis of *Streptomyces avermitilis* genes required for avermectin biosynthesis utilizing a novel integration vector." Gene **111**: 61-68.
- Malpartida, F. and D.A. Hopwood** (1986). "Physical and genetic characterisation of the gene cluster for the antibiotic actinorhodin in *Streptomyces coelicolor* A3(2)." Mol Gen Genet **205(1)**: 66-73.
- Maniatis, T., E.F. Fritsch, et al.** (1982). "Molecular cloning: a laboratory manual." Cold Spring Harbor, N.Y., Cold Spring Harbor Laboratory Press.
- Marck, C.** (1988). "DNA Strider: a C program for the fast analysis of DNA and protein sequence on the apple Macintosh family of computers." Nucleic. Acids. Res. **40**: 115-122.
- Marsden, A.F.A., P. Caffrey, et al.** (1994). "Stereo-specific acyl transfers on the erythromycin-producing polyketide synthase." Science **263**: 378-380.
- Martín, J. F. and L. E. McDaniel** (1977). "Production of polyene macrolide antibiotics." Adv. Appl. Microbiol. **21**: 1-52.
- Martín J.F. and P. Liras** (1989). "Organization and expression of genes involved in the biosynthesis of antibiotics and other secondary metabolites." Annu. Rev. Microbiol. **43**: 173-206.
- Maxam A.M. and W. Gilbert** (1980). "Sequencing end-labelled DNA with base specific chemical cleavages." Methods Enzymol. **65**: 449-560.
- Mazodier, P., R. Petter, et al.** (1989). "Intergeneric conjugation between *Escherichia coli* and *Streptomyces* species." J. Bacteriol **171**: 3583-3585.
- Mazodier, P., C.J. Thompson, et al.** (1990). "The chromosomal integration site of the *Streptomyces* element pSAM2 overlaps a putative tRNA gene conserved among actinomycetes." Mol. Gen. Genet. **222**: 431-434.

- McDaniel, R., S. Ebertkhosla, et al.** (1993a). "Engineered biosynthesis of novel polyketides: manipulation and analysis of an aromatic polyketide synthase with unproven catalytic specificities." J. Am. Chem. Soc. **115**: 11671-11675.
- McDaniel, R., S. Ebertkhosla, et al.** (1993b). "Engineered biosynthesis of novel polyketides." Sciences **262**: 1546-1550.
- Mediavilla A.** 1997. "Antibióticos macrólidos." *Farmacología humana*. 3ª edición, pp.: 1123-1130. Eds. Flórez J., Armijo J.A., Mediavilla A., Masson S.A., Barcelona.
- Méndez C. and J.A. Salas** (1998). "ABC transporters in antibiotic-producing actinomycetes." FEMS Microbiol. Lett. **158(1)**: 1-8. Review.
- Miller, J.H.** (1972). "Experiments in Molecular Genetics". Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
- Mólnar, I. J.F. Aparicio, et al.** (1996). "Organisation of the biosynthetic gene cluster for rapamycin in *Streptomyces hygroscopicus*: analysis of genes flanking the polyketide synthase." Gene **169**: 1-7.
- Moore, B.S. and J.N. Hopke** (2001). "Discovery of a new bacterial polyketide biosynthetic pathway." Chem. Biochem. **2(1)**: 35-38.
- Morosoli R., J.L. Bertrand, et al.** (1986). "Purification and properties of a xylanase from *Streptomyces lividans*." Biochem. J. **239**: 587-592.
- Motamedi H. and A. Shafiee** (1998). "The biosynthetic gene cluster for the macrolactone ring of the immunosuppressant FK506." Eur. J. Biochem. **256(3)**: 528-534.
- Mulks, M.H., M.G. Nair, et al.** (1990). "In vitro antibacterial activity of faerifungin, a new broad-spectrum polyene antibiotic macrolide." Antimicrob. Agents Chemoter **(34)9**:1762-1765.
- Muth, G., B. Nubbaumer, et al.** (1989). "A vector system with temperature-sensitive replication for gene disruption and mutational cloning in streptomycetes." Mol Gen Genet **219**: 341-348.
- Nair, M.G., A.R. Putnam, et al.** (1989). "Faerifungin: a new broad-spectrum antibiotic from *Streptomyces griseus* var. *autotrophicus*." J. Nat. Prod. **52**: 797-809.
- Ohnuki, T., T. Katoh, et al.** (1985). "Molecular cloning of tetracycline resistance genes from *Streptomyces rimosus* in *Streptomyces griseus* and characterization of the cloned genes." J. Bacteriol. **161(3)**: 1010-6.

- O'Keefe, D.P. and P.A. Harder** (1991). "Occurrence and biological function of cytochrome-P450 monooxygenases in the actinomycetes." Mol. Microbiol. **5**: 2099-2105.
- Olano, C., A.M. Rodríguez, et al.** (1995). "A second ABC transporter is involved in oleandomycin resistance and its secretion by *Streptomyces antibioticus*." Mol. Microbiol. **16(2)**: 333-343.
- Oliynyk, M., M.J. Brown, et al.** (1996). "A hybrid modular polyketide synthase obtained by domain swapping." Chem. Biol. **3**: 833-839.
- Omura, S. and H. Tanaka** (1984). "Production, structure and antifungal activity of polyene macrolides." *Macrolide Antibiotics, chemistry, biology and practice*. S. Omura. London, Academic Press: 351-404.
- Omura, S., H. Ikeda, et al.** (2001). "Genome sequence of an industrial microorganism *Streptomyces avermitilis*: Deducing the ability of producing secondary metabolites." Mol. Microbiol. **21**: 12215-12220.
- Onaka, H., N. Ando, et al.** (1995). "Cloning and characterization of the A-factor receptor gene from *Streptomyces griseus*." J. Bacteriol. **177**: 6083-6092.
- Paget, M.S.B., L. Chamberlin, et al.** (1999). "Evidence that the extracytoplasmic function sigma factor σ^E is required for normal cell wall structure in *Streptomyces coelicolor* A3(2)." J. Bacteriol. **181**: 204-211.
- Panzda K., G. Pfalzer, et al.** (1997). " Physical mapping shows that the unstable oxytetracycline gene cluster of *Streptomyces rimosus* lies close to one end of the linear chromosome." Appl. Environ. Microbiol. **143**: 1493-1501.
- Pigac J. and H. Schrempf.** (1995). " A simple and rapid method of transformation of *Streptomyces rimosus* R6 and other streptomycetes by electroporation." Appl. Environ. Microbiol. **61**: 352-356.
- Ramos, J.L., M. Martínez-Bueno, et al.** (2005). "The TetR family of Transcriptional Repressors." Microbiol. Mol. Biol. Rev. **69(2)**: 326-356.
- Redenbach, M., H.M. Kieser, et al.** (1996). "A set of ordered cosmids and a detailed genetic and physical map for the 8 Mb *Streptomyces coelicolor* A3(2) chromosome." Mol. Microbiol. **21(1)**: 77-96.

- Robinson, M., E. Lewis, et al.** (1981). "Occurrence of reiterated DNA sequences in strains of *Streptomyces* produced by an interspecific protoplast fusion." Mol. Gen Genet. **182(2)**: 336-340.
- Rodríguez, A.M., C. Olano, et al.** (1993). "*Streptomyces antibioticus* contains at least three oleandomycin resistance determinants, one of which shows similarity with proteins of the ABC transporter superfamily." Mol. Microbiol. **8**: 571-582.
- Rodríguez, A.M., C. Olano, et al.** (1995). "A cytochrome p450-like gene possibly involved in oleandomycin biosynthesis by *Streptomyces antibioticus*." FEMS Microbiol. Lett. **127**: 117-120.
- Ruan, X., D. Stassi, et al.** (1997). "A second type-I PKS gene cluster isolated from *Streptomyces hygroscopicus* ATCC29253, a rapamycin-producing strain." Gene **203(1)**: 1-9.
- Sambrook, J., E.F. Fritsch, et al.** (1989). "Molecular cloning: a laboratory manual." Cold Spring Harbor, NY, Cold Spring Harbor Laboratory Press.
- Sambrook, J. and D.W. Russell** (2001). "Molecular cloning: A laboratory manual." Cold Spring Harbor, NY, Cold Spring Harbor Laboratory Press.
- Sanger, F., S. Nicklen, et al.** (1977). "DNA sequencing with chain-terminating inhibitors." Proc. Natl. Acad. Sci. U.S.A. **74**: 5463-5467.
- Saxena, P., G. Yadav, et al.** (2003). "A new family of type III polyketide synthase in *Mycobacterium tuberculosis*." J. Biol. Chem. **278(45)**: 44780-44790.
- Schegel, R. and H. Thrum** (1968). "Flavomycoin, a new antifungal polyene antibiotic." Experientia **15**: 11-12.
- Schmitt-John T. and J.W. Engels** (1992). "Promoter constructions for efficient secretion expression in *Streptomyces lividans*." Appl. Microbiol. Biotechnol., **36**: 493-498.
- Schwecke, T., J.F. Aparicio, et al.** (1995). "The biosynthetic gene cluster for the polyketide immunosuppressant rapamycin." Biochemistry **92**: 7839-7843.
- Scrutton N.S., A. Berry, et al.** (1990). "Redesign of the coenzyme specificity of a dehydrogenase by protein engineering." Nature **343**: 38-43.
- Seco, E., F.J., Pérez Zúñiga, et al.** (2004). "Starter Unit Choice Determines the Production of Two Tetraene Macrolides, Rimocidin and CE-108. in *Streptomyces diastaticus* var., 108". Chem. and Biol. **11**: 357-366.

- Serre, L., E.C. Verbree, et al.** (1995). "The *Escherichia coli* malonyl-CoA:acyl carrier protein transacylases at 1.5-Å resolution. Crystal structure of a fatty acid synthase component." J. Biol. Chem. **270**: 12961-12964.
- Shrempf, H.** (1985). "Genetic instability: amplification, deletion and rearrangement within *Streptomyces* DNA." *Microbiology*-1985. American Society for Microbiology.
- Simon, R., U. Priefer, et al.** (1983). "A broad host range mobilisation system for in vivo genetic engineering: transposon mutagenesis in Gram-negative bacteria." Bio/Technology. **1**: 784-791.
- Southern, E.** (1975). "Detection of specific sequences among DNA fragments separated by gel electrophoresis." J. Mol. Biol. **98**: 503-517.
- Stackebrandt, E., W. Liesack, et al.** (1992). "Ribosomal RNA and rDNA Sequence Analyses." Gene **115(1-2)**: 255-260.
- Stassi, D.L., S. Donadio, et al.** (1993). "Identification of a *Saccharopolyspora erythraea* gene required for the final hydroxylation step in erythromycin biosynthesis." J. Bacteriol. **175**: 182-189.
- Stassi, D.L., S.J. Kakavas, et al.** (1998). "Ethyl-substituted erythromycin derivatives produced by directed metabolic engineering." Biochemistry **95**: 7305-7309.
- Stratigopoulos, G. and E. Cundliffe** (2002). "Expression analysis of the tylosin-biosynthetic gene cluster: pivotal regulatory role of the *tylQ* gene product." Chem. Biol. **9**: 71-78.
- Thompson, C.J., J.M. Ward, et al.** (1980). "DNA cloning in *Streptomyces*: resistance genes from antibiotic-producing species". Nature **286**: 525-527.
- Tohyama, H., Y. Okami, et al.** (1987). "Nucleotide sequence of the streptomycinphosphotransferase and amidinotransferase genes from *Streptomyces griseus*." Nucleic Acids Res. **15(4)**: 1819-1833.
- Vetlugina, L.A.** (1968). "The antibiotic roseofungin, its isolation and physico-chemical properties." Antibiotiki **13(11)**: 992-997.
- Vining, L.C.** (1992). "Secondary metabolism, inventive evolution and biochemical diversity: a review." Gene **115**: 135-140.
- Wakil, S.J.** (1989). "Fatty acid synthase, a proficient multifunctional enzyme". Biochemistry **28**: 4523-4530.

Walker, J.E., M. Saraste, et al. (1982). "Distantly related sequences in the α - and β - subunits ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold." EMBO J. **1**: 945-951.

Ward, J.M., G.R. Janssen, et al. (1986). "Construction and characterisation of a series of multi-copy promoter-probe plasmid vectors for *Streptomyces* using the aminoglycoside phosphotransferase gene from Tn5 as indicator." Mol Gen Genet **203(3)**: 468-478.

Wasserman, H.H., J.E. Vanverth, et al. (1967). "The mycoticins, polyene macrolides from *Streptomyces ruber*." J. Am. Chem. Soc. **89**: 1535-1536.

Weber, J.M., J.O. Leung, et al. (1991). "An erythromycin derivative produced by targeted gene disruption in *Saccharopolyspora erythraea*." Science **252**: 114-117.

Witkowski, A., A.K. Joshi, et al. (1999). "Conversion of a ketoacyl-synthase to a malonyl decarboxylase by replacement of the active-site cysteine with glutamine." Biochemistry **38**: 11643-11650.

Woese, D. (1987). "Bacterial evolution." Microbiol. Rev. **51**: 221-271.

Wright, F. and M.J. Bibb (1992). "Codon usage in the G+C-rich *Streptomyces* genome." Gene **113**: 55-65.

Wright, L.F. and D.A. Hopwood (1976). "Actinorhodin is a chromosomally- determined antibiotic in *Streptomyces coelicolor* A3(2)". J. Gen. Microbiol. **96(2)**: 289-297.

Yanish-Perron, C., J. Viera, et al. (1985). "Improved M13 phage cloning vector and host strains: nucleotide sequences of M13mp18 and pUC19 vectors." Gene **33**: 103-119.