



UNIVERSIDAD MICHOCANA DE SAN NICOLÁS DE HIDALGO

INSTITUTO DE INVESTIGACIONES
QUÍMICO-BIOLÓGICAS

“Evolución dirigida de la enzima 3-Metil- Crotonil-Coenzima A (MCCasa) carboxilasa de *Pseudomonas aeruginosa*”

Para obtener el Grado de:

DOCTOR EN CIENCIAS

Especialidad en:

Biología Experimental

Presenta:

MC. César Díaz Pérez

Asesor: Dr. Jesús Campos García

Coasesor: Dr. José Salud Rodríguez Zavala

ESTE TRABAJO SE REALIZÓ EN EL LABORATORIO DE
BIOTECNOLOGÍA MICROBIANA
DEL
INSTITUTO DE INVESTIGACIONES QUÍMICO BIOLÓGICAS
DE LA
UNIVERSIDAD MICHOACANA DE SAN NICOLÁS DE HIDALGO



DEDICATORIA

Dedico este trabajo a mi familia.

Mi padre y mi madre, José Pedro Díaz Sánchez y Estela Pérez Rodríguez†, ejemplo de trabajo y dedicación durante toda una vida, siempre me han alentando a dar lo mejor y ser un mejor ser humano.

Mis hermanos, Pedro, Claudia, Alma y Hugo, gracias por su comprensión y apoyo, me dieron fuerza en los momentos más difíciles.

Sin su incondicional apoyo nunca hubiera podido realizar mis metas.

He andado muchos caminos,
he abierto muchas veredas,
he navegado en cien mares
y atracado en cien riberas.

En todas partes he visto
caravanas de tristeza,
soberbios y melancólicos
borrachos de sombra negra,

y pedantones al paño
que miran, callan y piensan
que saben, porque no beben
el vino de las tabernas.

Mala gente que camina
y vaapestando la tierra...

Y en todas partes he visto
gentes que danzan o juegan
cuando pueden, y laboran
sus cuatro palmos de tierra.

Nunca, si llegan a un sitio,
preguntan adónde llegan.
Cuando caminan, cabalgan
a lomos de mula vieja,

y no conocen la prisa
ni aun en los días de fiesta.
Donde hay vino, beben vino;
donde no hay vino, agua fresca

Son buenas gentes que viven,
laboran, pasan y sueñan,
y en un día como tantos
descansan bajo la tierra.

"He andado muchos caminos"
Antonio Machado

AGRADECIMIENTOS

AL DC JESÚS CAMPOS GARCÍA, POR ACEPTARME EN SU GRUPO DE TRABAJO Y PERMITIRME LLEVAR A CABO ESTE TRABAJO. GRACIAS POR LOS DIFERENTES PUNTOS DE VISTA QUE ME AYUDARON A ENRIQUECERME Y DESARROLLARME EN IDEAS Y CONCEPTOS.

AL DC CARLOS CERVANTES VEGA, POR TODAS LAS SUGERENCIAS, IDEAS Y CONSEJOS.

A MIS SINODALES, DC CARLOS SOSA Y DC GUSTAVO SANTOYO, POR SUS OBSERVACIONES Y AYUDA EN LA REALIZACIÓN DE ESTA TESIS.

A LOS DC VÍCTOR MEZA CARMEN Y AGUSTINO MARTÍNEZ ANTONIO, POR SU TIEMPO Y LAS VALIOSAS OBSERVACIONES DURANTE LA REVISIÓN DE ESTA TESIS.

A MIS BUENOS AMIGOS MAU, LIZ, HANS, ESTHER, ARA, AMADA POR TODOS LOS MOMENTOS QUE ME AYUDARON A SEGUIR ADELANTE.

GRACIAS PAT, POR LAS BUENAS Y MALAS VIVENCIAS, LOS CONSEJOS Y OPINIONES, DISCUSIONES Y DEBATES QUE HEMOS TENIDO TODOS ESTOS AÑOS.

GRACIAS A MIS COMPAÑEROS LUZMA, AREPA, ANA, XOCH, CÉSAR, AK, ARNOLD, LORE, POR AGUANTARME SIEMPRE QUE LOS MOLESTABA.

A TODOS MIS COMPAÑEROS DE LABORATORIO DE BIOTECNOLOGÍA MICROBIANA, DE MICROBIOLOGÍA, DE FARMACOLOGÍA, IMPOSIBLE NOMBRARLOS A TODOS, PERO GRACIAS POR SU AYUDA, APOYO Y REACTIVOS SIN LOS QUE ESTE TRABAJO NO SE HUBIERA TERMINADO.

A MIS GRANDES AMIGOS DE IRAPUATO, LENIN, GLORIA, MARCH, IVÁN, LUIS DAVID, GINA, NIDIA, MARCE. ELLOS QUE SIEMPRE HAN CREÍDO EN MÍ A PESAR DE TODO; ELLOS QUE MARCARON UNA DE LA MEJORES ETAPAS DE MI VIDA; GRACIAS A ELLOS QUE ME HAN DADO UNA MANERA DISTINTA DE VER LA VIDA.

ÍNDICE

I) Resumen-----	v
II) Abstract-----	vii
III) Introducción-----	1
III.1) Biodegradación de hidrocarburos -----	1
III.1.a) Degradación de alcanos no recalcitrantes por microorganismos ---	1
III.1.b) Biodegradación de monoterpenos acíclicos de la familia del citronelol (MAFC)-----	2
III.1.c) Ruta de degradación del citronelol e interconexión con el catabolismo de leucina en <i>Pseudomonas aeruginosa</i> -----	3
III.2) Carboxilasas dependientes de biotina en las ruta MAFC-----	4
III.3)La 3-metilcrotonil Coenzima A Carboxilasa-----	6
III.4) Sitios activos de las CDB-----	8
IV) Justificación-----	9
V) Hipótesis-----	11
VI) Objetivos-----	11
VI.1) Objetivo general-----	11
VI.2) Objetivos específicos-----	11
VII) Estrategia experimental general-----	12
VIII) Resultados -----	13
VIII.1) Capitulo 1: “Análisis filogenético de la superfamilia de transcarboxilasas dependientes de coenzima A (TDC)” -----	13
VIII.2) Capitulo 2: “Co-expression of α and β subunits of the 3-Methyl- crotonyl-Coenzyme A carboxylase from <i>Pseudomonas aeruginosa</i> ”-----	41
VIII.3) Capitulo 3: “Structure modeling, docking and site-directed mutagenesis to identify residues of the active site of the β -subunit of 3- Methylcrotonyl-Coenzyme A carboxylase from <i>P. aeruginosa</i> : residues Ser187 and Arg51 are involved for substrate recognition in the MCCase”-----	52
IX) Discusión general -----	84

IX.1) Historia evolutiva de las transcarboxilasas dependientes de coenzima-A (TDC) -----	84
IX.2) Caracterización bioquímica de la MCCasa de <i>P. aeruginosa</i> -----	86
IX.3) Estado de oligomerización -----	86
IX.4) Modelaje de la MCCasa de <i>P. aeruginosa</i> -----	87
IX.5) Sitio activo de la MCCasa de <i>P. aeruginosa</i> -----	88
IX.6) Papel de los residuos R51, S187, F417, Y422 y G423 en el sitio activo de la MCCasa de <i>P. aeruginosa</i> -----	89
X) Conclusiones -----	92
XI) Perspectivas -----	93
XI.1) El complejo MCCasa -----	93
XI.2) La subunidad β de la MCCasa -----	94
XI.3) La subunidad α de la MCCasa -----	94
XI.4) El complejo GCCasa -----	95
XI.5) La familia de CDB -----	95
Referencias -----	96
Anexo 1 -----	101

ÍNDICE DE FIGURAS

Introducción

Figura 1.- Etapas generales de la degradación de los MAFC en <i>P. aeruginosa</i> -----	4
Figura 2.- Componentes centrales de la ruta catabólica ATU y LIU en <i>P. aeruginosa</i> -----	5
Figura 3.- Dominios que conforman a las carboxilasas dependientes de biotina (CDBs) que reconocen acil-CoA-----	7
Figura 4.- Sitios catalíticos de las CDB-----	9

Capítulo I

Figura 1.- Diagrama de los componentes de las tres clases de las CDB-----	18
Figura 2.- Diagrama del ordenamiento de genes y dominios de las carboxilasas dependientes de biotina-----	24
Figura 3.- Histograma de la longitud de las proteínas de la superfamilia de transcaboxilasas dependientes de CoA-----	26
Figura 4.- Topología del árbol filogenético de los subdominios a y b de la superfamilia de TDC-----	28
Figura 5.- Árbol filogenético de los subdominios a y b de la superfamilia de TDC-----	29
Figura 6.- Modelo evolutivo de la funcionalidad en la superfamilia TDC-----	35

Capítulo II

Figure 1.- Role of the 3-methylcrotonyl-CoA carboxylase from <i>P. aeruginosa</i> on the leucine/isovalerate and acyclic terpene pathways-----	45
Figure 2.- Purification and western-blot analysis of the co-expressed and co-purified MCCase from <i>P. aeruginosa</i> -----	48
Figure 3.- Determination of the molecular weight and oligomerization state of the MCCase from <i>P. aeruginosa</i> .-----	48
Figure 4.- Kinetic behavior of the co-expressed and co-purified MCCase from <i>P. aeruginosa</i> -----	49
Figure S1.- A) Co-expression plasmid pCDB3-----	51

Capítulo III

Figure 1.- Metabolic role and subunits of the 3-methyl-crotonyl-CoA carboxylase from <i>P. aeruginosa</i> PAO1-----	57
Figure 2.- Homology model of the MCCase β -subunit from <i>P. aeruginosa</i> --	63
Figure 3.- Insertions of amino acids in subdomains of BDC prevent the formation of large complexes-----	64
Figure 4.- Docking analysis of the substrate MC-CoA in the active site of the MCCase from <i>P. aeruginosa</i> -----	66
Figure 5.- Sequence alignment of the zones I and II of the MC-CoA-binding site of MCCase-----	68
Figure 6.- Docking analysis of the mutants of MCCase of <i>P. aeruginosa</i> ----	71

Figure S1.- Purification of the co-expressed and co-purified MCCase enzyme from <i>P. aeruginosa</i> -----	78
Figure S2.- Structural representation of the MCCase-----	78
Figure S3.- Ramachandran plot of modeled MCCase-----	79
Figure S4.- ProSA energy profiles-----	80
Figure S5.- Verified-3D analysis-----	81
Figura V1.- Diagrama de flujo para la obtención de los modelos de la subunidad β de la MCCasa de <i>P. aeruginosa</i> -----	82
 Discusión	
Figura 1.- Árboles filogenéticos de las subunidades α y β de las CDB-----	85
Figura 2.- Topología de la subunidad β de la MCCasa de <i>P. aeruginosa</i> ----	88
Figura 3.- Localización del sitio activo de la MCCasa de <i>P. aeruginosa</i> -----	90

I) RESUMEN

Las enzimas dependientes de biotina forman una superfamilia ubicua en los tres dominios de la vida, catalizan la carboxilación o descarboxilación de una amplia variedad de sustratos y son enzimas clave en las rutas metabólicas donde participan. Una de las enzimas de esta superfamilia, la 3-metilcrotonil-CoA carboxilasa (MCCasa; E.C.6.4.1.4), cataliza la carboxilación del 3-metilcrotonil-CoA (MCCoA) para formar 3-metilglutaconil-CoA (MGCoA). La MCCasa de la bacteria Gram negativa *Pseudomonas aeruginosa* ha sido objeto de estudio en nuestro grupo de trabajo y se le ha involucrado en el catabolismo de leucina y de monoterpenos acíclicos de la familia del citronelol.

Las enzimas de esta superfamilia están formadas por varios dominios proteicos bien definidos, los cuales son esenciales para llevar a cabo su función. El dominio de transcarboxilación (CT) se encarga del reconocimiento del sustrato a carboxilar y de la transferencia del grupo carboxilo. En este trabajo se reconstruyó la historia evolutiva de los dominios CT, encontrándose que estos dominios forman una superfamilia que reconoce sustratos con coenzima A, por lo que se nombró como superfamilia de transcarboxilasas dependientes de CoA (TDC). Esta superfamilia está formada por dos linajes, el primero que corresponde a las proteínas que codifican el dominio en dos genes (sCT) y el linaje de las proteínas que codifican al dominio CT en un solo gen (cCT). Para entender la historia evolutiva de los dominios se efectuó la filogenia de los subdominios del dominio CT, la cual arroja nueve familias, que se dividen de acuerdo al sustrato que reconocen. Los árboles filogenéticos indican que todos los miembros de la familia están formados por dos subdominios, los cuales probablemente surgieron de un evento de duplicación génica.

Al analizar la estructura de los miembros cristalizados de la superfamilia TDC, se observó que la unidad funcional es un dímero de dominios CT. Cada unidad funcional forma dos sitios activos, localizados en las zonas de interacción del dímero. Se realizó el modelaje del dímero de dominios CT de la MCCasa de *P. aeruginosa* y

el anclaje molecular del sustrato MCCoA en el modelo obtenido. El sitio activo está formado por la zona I que une y dirige al sustrato y la zona II donde se lleva a cabo la carboxilación de este. En esta zona también interacciona el cofactor biotina, que lleva el grupo carboxilo que se transferirá al sustrato.

Por otra parte, se llevó a cabo la expresión de los genes que codifican a la MCCasa en un vector de co-expresión, lo cual permitió implementar una estrategia rápida y eficiente para obtener la enzima y realizar su caracterización bioquímica. La MCCasa co-expresada presentó un complejo activo $\alpha_4\beta_4$, una K_m de 168 μM para el MC-CoA y una V_{max} de 430 $\text{nmol}\cdot\text{mg}^{-1}\cdot\text{min}^{-1}$, siguiendo una cinética hiperbólica.

Los análisis bioinformáticos realizados sugirieron que los residuos de la zona I, arginina 51 y serina 187, interaccionan con el sustrato para dirigirlo y fijarlo al sitio activo. En la zona II el análisis sugirió que la fenilalanina 417, la glicina 423 y la tirosina 422, podrían ser importantes para el reconocimiento del sustrato, dado que son residuos cercanos a la zona de carboxilación y son variables en la enzima paróloga geranil-CoA carboxilasa que también reconoce MC-CoA. Mediante mutagénesis sitio dirigida se probó tal hipótesis. En la zona I, la mutación R51A resultó en una disminución del 50% en la actividad de la MCCasa. En la zona II el cambio F417Y aumentó casi 3 veces la actividad de la enzima y por el contrario la mutación G423A presentó una disminución del 30% en la velocidad de reacción. El cambio en las constantes cinéticas de la enzima sugiere que estos residuos se encuentran en el sitio activo de la MCCasa y están involucrados en la interacción con el sustrato. Por otro lado, las mutaciones S187A y Y422D abatieron por completo la actividad de la MCCasa, por lo que se concluye que estos residuos son esenciales para la actividad enzimática.

II) ABSTRACT

Biotin dependent enzymes form an ubiquitous superfamily in the three domains of life; these enzymes catalyze the carboxylation or decarboxylation of a wide variety of substrates, as well as they play a key role in the metabolic pathways where these proteins participate. One enzyme of this superfamily is the 3-methylcrotonyl-CoA carboxylase (MCCase; E.C.6.4.1.4), which catalyze the 3-methylcrotonyl-CoA (MCCoA) carboxylation to produce 3-methylglutaconyl-CoA (MGCoA). The MCCase from *Pseudomonas aeruginosa* has been studied in our group, and we have seen that it is involved in the catabolism of leucine and acyclic monoterpenes of citronellol family.

Enzymes of this superfamily are composed by several protein domains, which are essential to carry out their function. The transcarboxylation domain (CT) is responsible for the recognition of the substrate to be carboxylated and for the transference of the carboxyl group. In this work, the evolutionary history of the CT domain that recognizes the substrates-CoA has been constructed. The findings indicate that these domains constitute a superfamily with diverse substrate-CoA recognition, named here as transcarboxylase CoA-dependent superfamily (TDC). The TDC superfamily is constituted by two lineages, the first one corresponds to proteins that encode the CT domain in two separated genes (sCT), and the second lineage contain proteins that the encode CT domain in a single gene (cCT). To understand the evolutionary history of these domains, a phylogenetic analysis of the subdomains from the CT-domain was conducted. Phylogenetic trees showed that the TDC superfamily is constituted by nine main groups, according to the recognized substrates. The analysis also shows that all members of this superfamily are formed by two subdomains, which probably derivate from a very ancient gene duplication event.

Structural analysis of crystallized members of TDC superfamily shows that the functional unit is a dimer of CT domains. Each functional unit contains two active sites, which are located at the dimer-interaction zone. Molecular modeling of CT

dimer domains from the *P. aeruginosa* MCCase was built. The model was used to perform docking analysis with MCCoA substrate. The active site is formed by two zones: zone I, involved in substrate recognition and binding, and zone II, where carboxylation of the MCCoA substrate is carried out. In zone II takes place the interaction with the biotin cofactor, which carries the carboxyl group that is transferred to the substrate.

Expression of MCCase genes in a co-expression vector allowed us to efficiently express and purify the MCCase to perform its biochemical characterization. The co-expressed MCCase showed an active complex $\alpha_4\beta_4$, with a K_m of 168 μM for MCCoA and V_{max} of 430 $\text{nmoles}\cdot\text{mg}^{-1}\cdot\text{min}^{-1}$, with a hyperbolic kinetics.

Bioinformatic analysis suggested that the residues located in the zone I, arginine 51 and serine 187, interact with the substrate to be coupled and conducted into the active site. In the zone II, the analysis suggested that the phenylalanine 417, glycine 423 and tyrosine 422 (neighbors residues located in the carboxylation zone and variable in the paralog enzyme geranyl-CoA carboxylase that also recognizes MCCoA), could be important for the substrate recognition. Therefore, site-directed mutagenesis was used to test this hypothesis. Mutation in the residue R51A showed a decreased MCCase activity, while a mutation in F417Y increased the activity of the enzyme and conversely the mutation G423A showed a decreasing in V_{max} . The changes in the kinetic constants of the enzyme suggested that these residues are located in the active site of the MCCase and these were involved the interaction with the substrate. On the other hand, mutants S187A and Y422D abolished completely the MCCase activity, thus, it is concluded that these residues are essential for the enzymatic activity of the MCCase protein.

III) INTRODUCCIÓN

La contaminación ambiental es uno de los grandes problemas a los que se enfrenta la humanidad hoy en día. En las últimas décadas se han liberado al ambiente una gran cantidad de compuestos químicos debido al aumento de la población, el uso desmedido de fertilizantes y plaguicidas agrícolas y a la expansión de la industria alimenticia y de manufactura (Díaz-Pérez 2003).

Los hidrocarburos son uno de los contaminantes ambientales de mayor importancia en el mundo. Se estima que anualmente son arrojados hasta 8.4 millones de toneladas al ambiente, lo cual conlleva a un gran daño ecológico (NCR 2003).

III.1) Biodegradación de hidrocarburos

En el ambiente las bacterias de vida libre frecuentemente cuentan con un metabolismo versátil que les permite usar una amplia gama de compuestos como fuentes de carbono y de energía (Rojo 2010). Algunas de estas bacterias se han vuelto especialistas como *Alcanivorax borkumensis*, una bacteria marina capaz de asimilar alcanos lineales y ramificados, sin embargo, es incapaz de metabolizar azúcares, aminoácidos, ácidos grasos e hidrocarburos aromáticos (Yakimov *et al.* 1998; Schneiker *et al.* 2006).

Los hidrocarburos son degradados por los sistemas biológicos predominantemente vía la oxidación inicial de un grupo metilo terminal, hasta un ácido carboxílico seguida por una secuencia de oxidación β , lo cual implica el acoplamiento de más de una ruta metabólica para la modificación de las moléculas en intermediarios asimilables (Pirnik *et al.* 1974; Fall *et al.* 1979; Campos-García 2000).

III.1.a) Degradación de alcanos no recalcitrantes por microorganismos

Los hidrocarburos saturados, conocidos comúnmente como alcanos, pueden ser lineales, cíclicos o ramificados, y constituyen hasta el 50% del contenido en el petróleo. Estos compuestos pueden ser degradados aeróbica o anaeróbicamente por bacterias (Rojo 2009). Cabe destacar que de acuerdo a la complejidad de su

estructura química, estos compuestos pueden ser asimilados rápidamente, o permanecer largos periodos en el ambiente, es decir ser recalcitrantes (Alexander 1973; Pirnik *et al.* 1974; Fall *et al.* 1979; Schaeffer *et al.* 1979).

Los alcanos lineales son degradados por un gran número de microorganismos, se han descrito más de 20 géneros bacterianos, 11 géneros de levaduras y 11 géneros de hongos filamentosos, por lo cual los alcanos lineales no persisten en el ambiente (Van Beilen *et al.* 2003; Wentzel *et al.* 2007). Uno de los sistemas degradativos de alcanos lineales más conocido es el codificado en el operón *alkBFGHJKL* del plásmido OCT, de *Pseudomonas putida* GPo1 (antes conocida como *Pseudomonas oleovorans*), en donde ácidos grasos de 6-12 átomos de carbono son transportados y activados con coenzima A por los productos génicos de este operón (van Beilen *et al.* 2001). Los acil-CoA posteriormente entran a oxidación β para ser asimilados o incorporados a rutas catabólicas centrales.

La degradación de los alcanos ramificados es más compleja, sin embargo, aquellos que presentan una ramificación en el carbono 2 (carbono α) son degradados mediante el acoplamiento de la oxidación α , la oxidación ω y la oxidación β . El ejemplo más estudiado es la degradación del pristano (2,6,10,14-tetrametil-pentadecano) por *Brevibacterium erythrogenes* (Pirnik *et al.* 1974).

III.1.b) Biodegradación de monoterpenos acíclicos de la familia del citronelol (MAFC)

La estructura química de los hidrocarburos afecta la velocidad en la que son degradados, observándose que los compuestos lineales son asimilados más rápidamente que los compuestos que presentan ramificaciones (Dias y Alexander 1971; Schaeffer *et al.* 1979). La familia de los hidrocarburos 3-metil ramificados presenta una ramificación sobre el carbono 3 ó carbono β de la cadena principal, bloqueando la oxidación β , por lo que son clasificados como recalcitrantes (Alexander 1973; Pirnik *et al.* 1974; Schaeffer *et al.* 1979). Los isoprenoides acíclicos de la familia del citronelol (MAFC) como el citronelol, geraniol, nerol, citronelal, citral,

ácido citronélico y ácido geránico pertenecen a este tipo de compuestos (Díaz-Pérez *et al.* 2004), de los cuales el citronelol es el modelo para estudiar la biodegradación por microorganismos (Fall *et al.* 1979; Schaeffer *et al.* 1979). Actualmente solo se conocen algunas especies del género *Pseudomonas* que son capaces de metabolizar al citronelol, mediante una ruta degradativa que fue propuesta en *Pseudomonas citronellolis* (Cantwell *et al.* 1978).

III.1.c) Degradación del citronelol e interconexión con el catabolismo de leucina en *Pseudomonas aeruginosa*

Pseudomonas aeruginosa es una bacteria Gram negativa versátil, capaz de crecer en un gran número de ambientes como suelos, costas, hábitats marinos, tejidos de plantas y animales (Stover *et al.* 2000). *P. aeruginosa* además puede utilizar al citronelol como única fuente de carbono y energía, por lo que se ha usado como modelo biológico en la búsqueda de los genes involucrados en la degradación de los MAFC (Díaz-Pérez *et al.* 2004; Hoschle *et al.* 2005).

La ruta de degradación de los MAFC comprende cuatro etapas: (I) la ruta superior de oxidación-activación, (II) ruta central de terpenos acíclicos (ATU), (III) acoplamiento con oxidación β y (IV) convergencia con la ruta de degradación de leucina/isovalerato (LIU) (Fig. 1) (Campos-García 2010).

Las rutas catabólicas centrales ATU y LIU son interesantes ya que ambas llevan a cabo un grupo de reacciones análogas que involucran a una carboxilasa dependiente de biotina (CDB) y a una hidratasa, que son específicas para cada intermediario, además probablemente comparte la liasa (LiuE) que es capaz de usar los substratos de ambas rutas (Fig. 2A). Más aún, se ha observado que la subunidad α de las carboxilasas son homólogas entre sí, y que las proteínas de las rutas tienen identidad de secuencia de aminoácidos (Fig. 2B), lo que sugiere que ambas rutas provienen de un origen común (Aguilar *et al.* 2006; Chavez-Aviles *et al.* 2010).

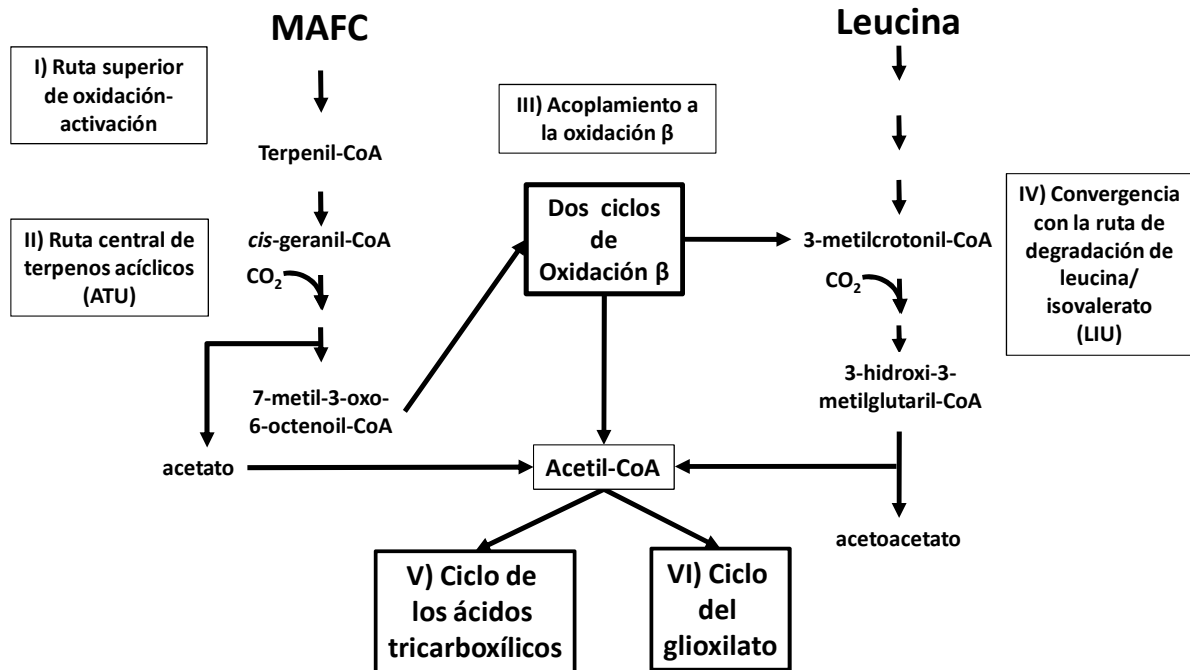


Figura 1.- Etapas generales de la degradación de los MAFC en *P. aeruginosa*. Los MAFC como el citronelol entran a la ruta superior donde se oxidan y se activan con Co-A (I). Una vez activados los metabolitos pasan a la ruta central ATU, donde la ramificación en la posición 3 es removida con la producción de acetato y 7-metil-3-oxo-6-octenoil-CoA (II), el cual puede entrar a la oxidación β (III), para producir el MC-CoA, que es el punto de convergencia con el catabolismo del aminoácido leucina e isovalerato, donde se produce acetoacetato y acetil-CoA (IV), el cual es canalizado al ciclo de los ácidos tricarboxílicos (V) o al ciclo del glioxilato (VI) (Adaptado de Campos-García 2010).

III.2) Carboxilasas dependientes de biotina en la ruta MAFC

Las carboxilasas dependientes de biotina (CDB) son una superfamilia de enzimas que derivan su nombre del hecho de que tienen como grupo prostético a la biotina (vitamina H). Están divididas en tres clases (I, II y III) de acuerdo al tipo de función que llevan. La clase I agrupa a las carboxilasas, la clase II incluye a las descarboxilasas y la clase III contiene a las transcarboxilasas (Hall *et al.* 2004). Estas enzimas llevan a cabo la transferencia de un grupo carboxilo a un substrato, con consumo de ATP. Se ha observado que las CDB son enzimas clave en las rutas metabólicas donde se presentan (Chapman-Smith *et al.* 2001; Jitrapakdee y Wallace 2003; Nikolau *et al.* 2003; Campos-García 2010). Las CDB pueden estar formadas por una o más cadenas polipeptídicas, por ejemplo, la acetil-CoA carboxilasa

(ACCasa) de eucariotes consta de una sola cadena de más de 2000 aminoácidos (aa), la ACCasa de bacterias está formada por cuatro cadenas polipeptídicas de diversos tamaños (Zhang *et al.* 2003), mientras que la propionil-CoA carboxilasa (PCCasa), la MCCasa y la geranil-CoA carboxilasa (GCCasa) de bacterias están compuestas por dos cadenas proteínicas (Fig. 3) (Jitrapakdee *et al.* 2003; Aguilar *et al.* 2006; Campos-García 2010; Huang *et al.* 2010).

Una característica común de todas las CDB es que contienen tres dominios indispensables para su función: (i) el dominio de carboxilación de biotina (BC); (ii) el dominio acarreador de carboxi-biotina (BCCP) y (iii) el dominio de transferencia de carboxilo (CT) (Jitrapakdee *et al.* 2003; Nikolau *et al.* 2003). A su vez, el dominio CT se compone de dos subdominios, a y b (Fig. 3) (Zhang *et al.* 2003).

En las rutas ATU y LIU existen dos enzimas clave: la geranil-CoA carboxilasa (GCCasa) y la 3-metil-crotonil-CoA carboxilasa (MCCasa). La GCCasa y la MCCasa se encuentran constituidas por dos subunidades cada una, denominadas α y β (Fig. 3). En ambos casos la subunidad α contiene los dominios BC y BCCP, mientras que la subunidad β contiene el dominio TC (Fig. 3). La GCCasa y la MCCasa se encuentran codificadas por los genes *atuC/atuF* y *liuB/liuD*, respectivamente (Fig. 2A) (Hoschle *et al.* 2005; Aguilar *et al.* 2006).

La reconstrucción de la historia evolutiva de la subunidad α de la GCCasa y MCCasa de *P. aeruginosa* sugiere que son proteínas homólogas, lo cual se ve reflejado en la identidad del 51% entre dichas subunidades; en el árbol filogenético de esta familia, las subunidades α de la MCCasa y de la GCCasa se ubican en la misma rama formando parte de un mismo grupo (Aguilar *et al.* 2006).

III.3) La 3-metilcrotonil Coenzima A carboxilasa

La enzima MCCasa, (E.C.6.4.1.4) cataliza la carboxilación del 3-metilcrotonil-CoA (MCCoA) para formar el 3-metil-glutaconil-CoA (MGCoA), en una reacción

dependiente de ATP que se lleva a cabo en dos pasos (Aguilar *et al.* 2006; Campos-García 2010).

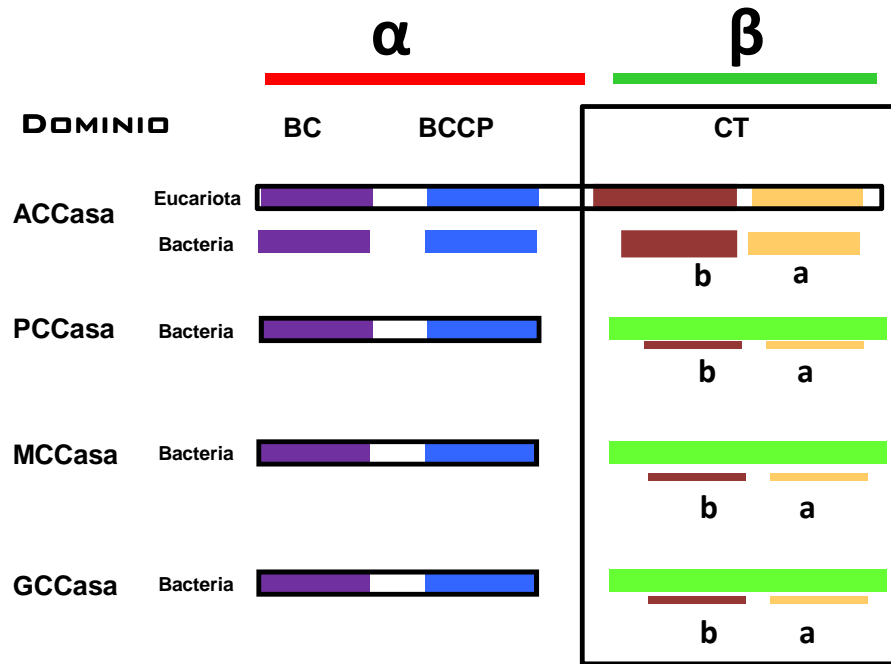
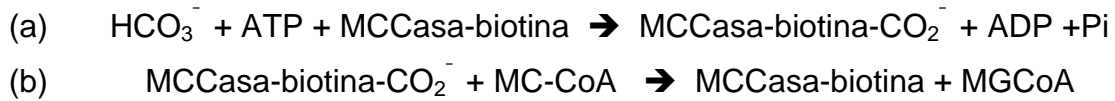


Figura 3.- Dominios que conforman a las carboxilasas dependientes de biotina (CDBs) que reconocen acil-CoA. En la figura se presentan los esquemas de las subunidades α y β de las carboxilasas y los dominios que conforman cada subunidad. Se representan las CDBs: acetil-CoA carboxilasa (ACCasa), propionil-CoA carboxilasa (PCCasa), 3-metil-crotonil-CoA carboxilasa (MCCasa) y geranil-CoA carboxilasa (GCCasa). En el recuadro se muestra el dominio CT que está formado por dos subdominios (a y b).

La MCCasa se ha localizado en bacterias (Fall 1981; Díaz-Pérez *et al.* 2004; Hoschle *et al.* 2005), hongos (Rodríguez *et al.* 2004), plantas (Alban *et al.* 1993; Wurtele y Nikolau 2000) y mamíferos (Lau *et al.* 1979; Chu y Cheng 2007). En humanos, las mutaciones en los genes que codifican para esta enzima provocan la enfermedad conocida como metilcrotonilglicinuria (Desviat *et al.* 2003).

En *P. aeruginosa* la MCCasa se expresa cuando se cultiva a la bacteria en medios con leucina o citronelol, debido a que su substrato es el punto de convergencia de la ruta ATU y LIU (Campos-García 2010). Su importancia se resalta, ya que cepas mutantes en genes de cualquiera de las subunidades de la MCCasa no son capaces de crecer en leucina o citronelol como única fuente de carbono (Aguilar *et al.* 2006).

Por otra parte, se ha observado que la MCCasa de *P. citronellolis* presenta una estructura cuaternaria $\alpha_4\beta_4$ (Fall y Hector 1977), mientras que en eucariotes se ha observado como $\alpha_4\beta_4$ y $\alpha_6\beta_6$ (Lau *et al.* 1980; Wurtele *et al.* 2000).

III.4) Sitios activos de las CDB

Las CDB presentan dos sitios activos, uno en el dominio BC que se encarga de la carboxilación de biotina (Sitio I), y el segundo en el dominio TC (Sitio II), donde se realiza la transferencia del grupo carboxilo desde la biotina al acil-CoA. Se ha observado que para que la transferencia pueda llevarse a cabo, el dominio BCCP debe translocarse para poder cubrir la distancia entre ambos sitios catalíticos (Fig. 4) (St Maurice *et al.* 2007; Huang *et al.* 2010).

Todas las subunidades α de la familia de las CDB reconocen los mismos substratos (ATP y bicarbonato) y llevan a cabo la misma reacción química (la carboxilación del cofactor biotina), por lo que no es de extrañar que en las enzimas de estructura conocida PCCasa de la proteobacteria *Silicibacter pomeroyi* y piruvato carboxilasa (PC) de la proteobacteria *Rhizobium etli* la localización y la conformación de los residuos que interactúan con los substratos en el sitio activo se encuentren conservados (Huang *et al.* 2010).

La subunidad β de la MCCasa presenta similitud de secuencia con la subunidad β de la PCC de *S. pomeroyi* y la subunidad 12S de la transcarboxilasa (MTCasa) de la actinobacteria *Propionibacterium freudenreichii*. Las PCCasas y la MTCasas presentan su sitio activo II en la interface de un dímero de subunidades β (Hall *et al.* 2003; Zhang *et al.* 2003; Diacovich *et al.* 2004; Huang *et al.* 2010). Sin embargo, en

la MCCasa de *P. aeruginosa*, se desconoce su estado de oligomerización, su sitio activo, y los aminoácidos que estén formando el hueco (pocket) catalítico.

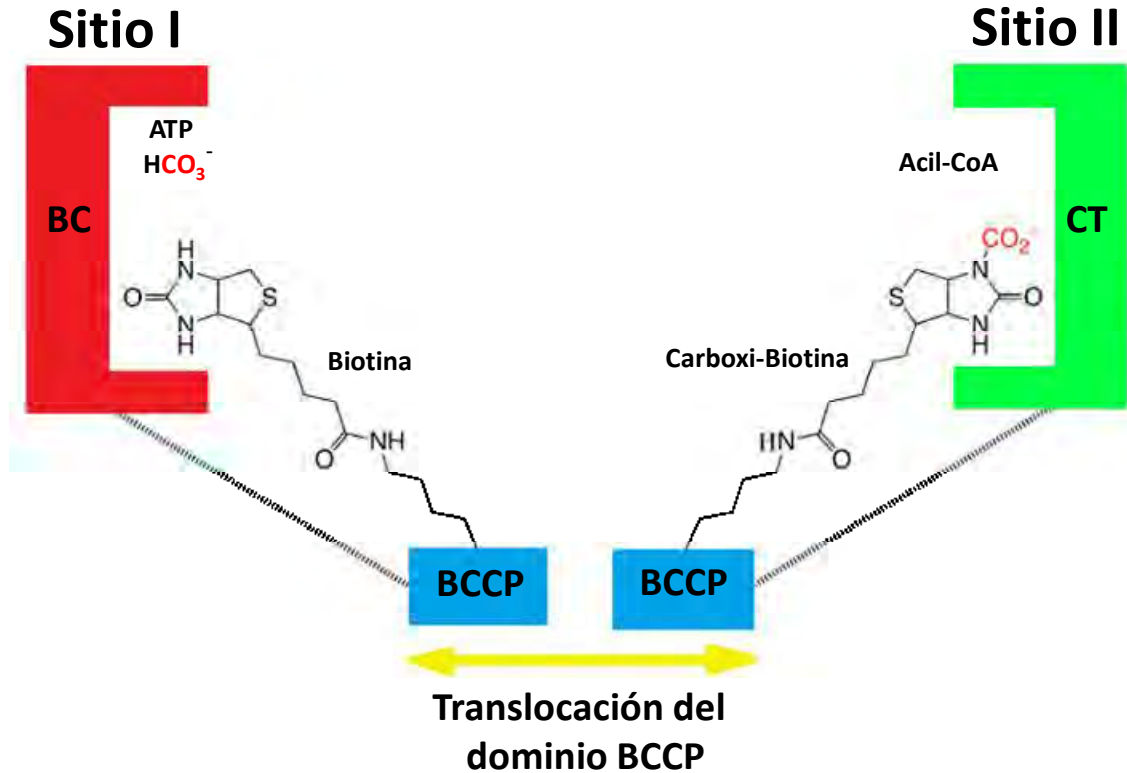


Figura 4.- Sitios catalíticos de las CDB. Todas las CDB presentan dos sitios catalíticos, localizados en los dominios BC y CT, respectivamente. El tercer dominio BCCP sirve como acarreador del cofactor biotina y de enlace para que se transfiera el grupo carboxil entre los dos sitios.

IV) JUSTIFICACIÓN

Si bien se conoce el sitio activo de diversas CDB, como en la PCCasa, ACCasa y TDCasa, en la MCCasa se desconoce su estructura tridimensional, así como la localización del sitio activo de transferencia del carboxilo y los aminoácidos que pueden estar conformándolo. En la subunidad β de la MCCasa de *P. aeruginosa* se localiza el dominio CT y posiblemente el sitio II, que es en donde se lleva a cabo el reconocimiento del sustrato, el MC-CoA. Es de interés investigar cuáles son los aminoácidos que forman parte del sitio activo y dilucidar la importancia de ellos.

En la subunidad β de la MCCasa y la GCCasa, se realiza la transferencia de un grupo carboxilo, sin embargo, el reconocimiento del sustrato es diferencial, aunado a que la identidad de secuencia ha disminuido a 46%, en comparación con el 51% de identidad que existe entre las subunidades α de estas CDBs. Esta diferencia llega a un 30% de identidad si se toma en cuenta a otras subunidades β de las familias de las CDB. Lo que lleva a preguntarnos si la relación evolutiva de la subunidad β es la misma que la de la subunidad α , o si ambas subunidades han seguido un camino divergente.

Los abordajes filogenéticos que se han llevado a cabo para entender a las CDB han sido limitados en el número de secuencias que se tienen, enfocados a un grupo de organismos cercanos taxonómicamente, y se ha tomado en cuenta a otros dominios (Toh *et al.* 1993; Tu y Hagedorn 1997; Jordan *et al.* 2003; Li *et al.* 2010), los cuales no tienen identidad de secuencia con el dominio CT de las carboxilasas que reconocen acil-CoA, lo cual puede bajar la calidad del alineamiento y de la reconstrucción, por lo que una filogenia masiva de los dominios CT de las carboxilasas de acil-CoA ayudaría a entender de mejor manera la evolución de estas enzimas.

V) HIPÓTESIS

Mutaciones en la subunidad β de la 3-metilcrotonil-coenzima A carboxilasa de *Pseudomonas aeruginosa* modificarán el reconocimiento por su substrato y por el geranil-CoA.

VI) OBJETIVOS

VI.1) Objetivo general

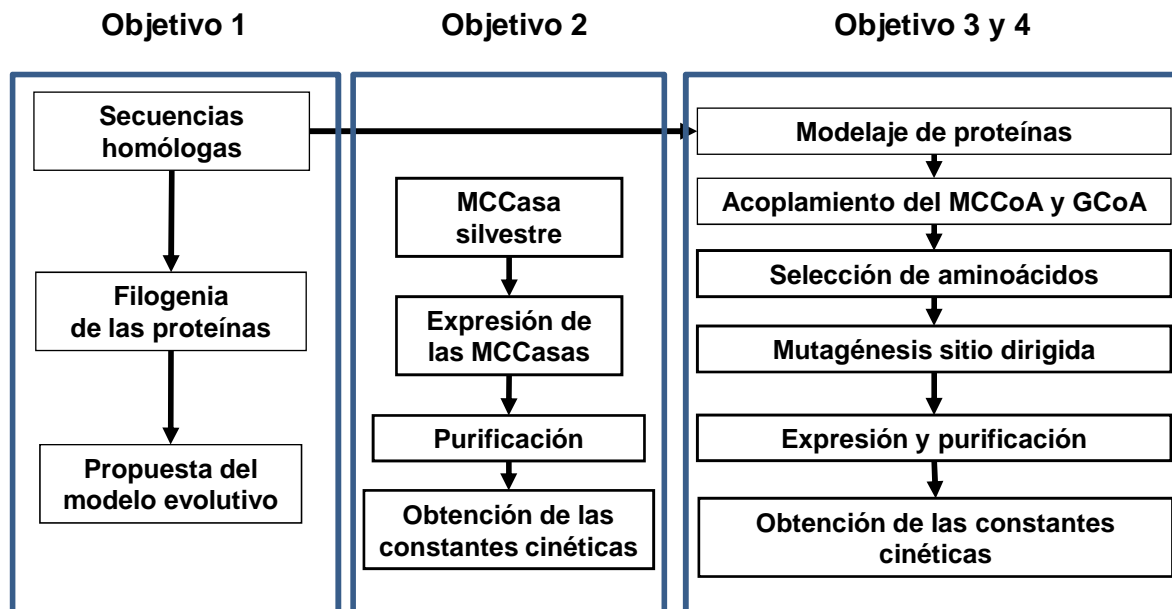
Dilucidar el papel de los aminoácidos del sitio activo de la enzima 3-metilcrotonil-coenzima A carboxilasa de *Pseudomonas aeruginosa* en el reconocimiento del substrato.

VI.2) Objetivos específicos

1. Construir un modelo de evolución de las enzimas transcarboxilasas dependientes de biotina.
2. Caracterizar bioquímicamente la enzima de 3-metilcrotonil-CoA carboxilasa de *Pseudomonas aeruginosa*.
3. Modelar *in silico* las enzimas 3-metilcrotonil-CoA carboxilasa y geranil-CoA carboxilasa de *Pseudomonas aeruginosa* y realizar ensayos de acoplamiento molecular con los substratos MC-CoA y G-CoA.
4. Caracterizar y discernir el papel de los aminoácidos involucrados en el reconocimiento del substrato de la 3-metilcrotonil-CoA carboxilasa de *Pseudomonas aeruginosa*.

VII) ESTRATEGIA EXPERIMENTAL GENERAL

La estrategia general se dividirá en tres partes. La primera involucra la reconstrucción de la historia evolutiva de la subunidad β de la MCCasa; la segunda consiste en la generación del modelo tridimensional de la subunidad β de la MCCasa en su forma funcional y el anclaje del sustrato en el modelo; y la tercera parte comprende la caracterización cinética de la MCCasa silvestre y discernir el papel de los aminoácidos involucrados en el reconocimiento del sustrato mediante mutagénesis sitio-dirigida.



VIII.1) Capítulo I:

“Análisis filogenético de la superfamilia de transcarboxilasas dependientes de Coenzima A (TDC)”

Díaz-Pérez César, Rodríguez-Zavala José Salud, Díaz-Pérez Alma-Laura y Campos-García Jesús.

VIII.1.a) RESUMEN

Las carboxilasas dependientes de biotina (CDB) son enzimas clave en una gran variedad de rutas metabólicas, como la biosíntesis y degradación de ácidos grasos, el catabolismo de aminoácidos como la leucina, la degradación de compuestos isoprenoides, entre otros. El mecanismo mediante el cual estas enzimas llevan a cabo su función es en dos reacciones: la carboxilación del cofactor biotina, seguida de la transferencia del grupo carboxilo hacia el sustrato. Cada reacción se lleva a cabo en dominios específicos. El dominio de carboxilación de la biotina es muy parecido en todas las CDB, sin embargo, el dominio de transcarboxilación es diferente, existiendo tres diferentes tipos, el CT1 que reconoce compuestos con CoA, el CT2 que reconoce piruvato, y el CT3 que reconoce urea. El dominio CT1 puede encontrarse como dominio único en una proteína, ó ser parte de una proteína multidominio, aunque también puede estar formado por dos proteínas, las cuales interactúan para formar el dominio funcional. El CT1 reconoce la variedad más grande de sustratos, por lo que es de nuestro interés investigar su historia evolutiva. En este trabajo se llevó a cabo la filogenia molecular de las proteínas que forman al dominio CT1. Se encontró que la superfamilia de transcarboxilasas CT1 consta de 2556 miembros. El dominio CT1 consta de dos subdominios, que pueden estar localizados en dos proteínas pequeñas que en pareja forman un dominio CT1 funcional (sCT), o puede encontrarse en una sola proteína que tiene los dos subdominios fusionados (cCT). Las enzimas de esta superfamilia reconocen sustratos con CoA, por lo que se le nombró superfamilia de transcarboxilasas dependientes de CoA (TDC). Al analizar los subdominios de las TDC se encontró que probablemente éstos surgieron de una duplicación de un dominio ancestral. La distribución filogenética indica la existencia de dos linajes principales, uno formado

por las sCT y otro por las cCT. A partir de estos dos linajes, los eventos evolutivos sugieren la formación de grupos separados que reconocen sustratos específicos.

VIII.1.b) ABSTRACT

The biotin-dependent carboxylases (BDC) are key enzymes in a variety of metabolic pathways, as biosynthesis and degradation of fatty acids, amino acids catabolism like leucine and isoprenoids degradation. The BDC mechanism is in two steps: the carboxylation of cofactor (biotin), followed by the transference of carboxyl group to the substrate. Each step is carried out in specific domains of the enzyme. The biotin carboxylase domain is very similar in all BDC's, however, the carboxy transferase domain (CT) is variable. Three different CT domains are known: the type CT1 recognizes compounds with CoA; the type CT2 uses pyruvate as substrate; and the type CT3 that recognizes urea. The CT1 domain can be found as unique domain in a protein, forming a multidomain protein or can be constituted by two separated peptides that interacts to form a functional domain. The CT1 domain recognizes a wide variety of substrates, so the objective of this work was to investigate its evolutionary history. In this work the molecular phylogeny of proteins that contain the CT1 was constructed. It was found that the CT1 domain proteins form a superfamily with 2556 members. The CT1 domain is constituted by two subdomains. The CT1 subdomains can be codified in two small proteins that form a CT1 functional domain (sCT) or can be codified in one large protein that contains both subdomains fused (cCT). All characterized members recognized CoA compounds, for this reason, the superfamily was named CoA-dependent carboxy transferases (CDT). It was found that the CT1 subdomains probable arisen from an ancient duplication from an ancestral single domain. The phylogenetic distribution suggests the existence of two principal lineages: one constituted by the sCT proteins, and other by the cCT enzymes. From these two lineages, the evolutive events suggest the formation of separated groups that recognizes specific substrates.

VIII.1.c) INTRODUCCIÓN

Las carboxilasas dependientes de biotina (CDB) forman una familia de enzimas que usan a la biotina (vitamina H) como cofactor, para llevar a cabo la transferencia de un grupo carboxilo a una diversa gama de sustratos (Jitrapakdee y Wallace 2003). Las enzimas de esta familia están involucradas en diversas rutas metabólicas donde juegan papeles clave o regulatorios (Tabla 1).

Tabla 1.- Rutas donde participan las enzimas CDB

Enzima	Ruta metabólica	Referencia
Acetil-CoA carboxilasa (ACCasa)	Biosíntesis ácidos grasos Oxidación de ácidos grasos	(Jitrapakdee <i>et al.</i> 2003)
Propionil-CoA carboxilasa (PCCasa)	Catabolismo de aminoácidos ramificados Catabolismo de ácidos grasos de cadena larga	(Jitrapakdee <i>et al.</i> 2003)
Metilcrotonil-CoA carboxilasa (MCCasa)	Catabolismo de leucina Degradación de monoterpenos acíclicos	(Aguilar <i>et al.</i> 2006; Aguilar <i>et al.</i> 2008)
Geranil-CoA carboxilasa (GCCasa)	Degradación de monoterpenos acíclicos	(Aguilar <i>et al.</i> 2006; Aguilar <i>et al.</i> 2008)
Urea-CoA carboxilasa (UCasa)	Degradación de urea	(Kanamori <i>et al.</i> 2004)
Glutaconil-CoA descarboxilasa (GDCasa)	Degradación de glutamato Degradación de glutaril-CoA	(Buckel 2001; Westover <i>et al.</i> 2001)
Metilmalonil-CoA descarboxilasa (MDCasa)	Vía de descarboxilación de succinato	(Haller <i>et al.</i> 2000)
Malonato descarboxilasa (MaDCasa)	Degradación de malonato	(Dimroth y Hilbi 1997)
Metilmalonil-CoA transcarboxilasa (MTCasa)	Metabolismo de propionato	(Houwen <i>et al.</i> 1990)
Piruvato carboxilasa (PCasa)	Gluconeogénesis Lipogénesis	(Jitrapakdee y Wallace 1999)
Oxalacetato descarboxilasa (ODCasa)	Metabolismo anaeróbico del citrato	(Dimroth <i>et al.</i> 2001)

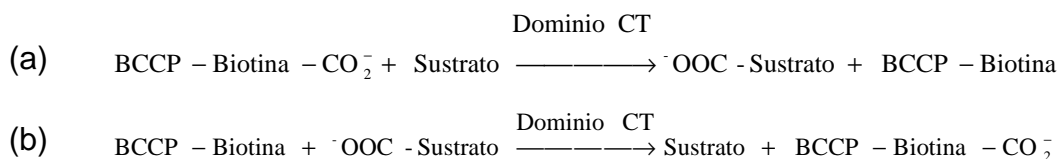
La familia de CDB puede ser dividida de acuerdo a la forma de transferencia del grupo carboxilo hacia el sustrato, así la clase I formada por las PC, ACCasa, PCCasa, MCCasa, GCCasa y UCasa, requieren ATP, HCO_3^- y Mg^{2+} para llevar a cabo su función; los miembros de la clase II, comprendida por las GDCasa, MDCasa, ODCasa, y MaDCasa, acoplan la descarboxilación del sustrato con el transporte de

sodio; y la clase III, formada por las MTCasas, llevan a cabo la descarboxilación del metilmalonil-CoA, acoplada a la carboxilación del piruvato (Hall *et al.* 2003).

Las CDB son una familia multidominio; las carboxilasas de la clase I cuentan con tres dominios: el dominio de carboxilación de biotina (BC), el dominio acarreador de biotina (BCCP), y el dominio de transcarboxilación (CT) (Jitrapakdee *et al.* 2003). Los miembros de la clase II presentan los dominios CT y BCCP, similares a los de la clase I, además cuentan con un dominio transmembranal (TM) (Buckel 2001). La clase III son formadas por un dominio BCCP, y por dos dominios CT (Figura 1) (Hall *et al.* 2003).

El dominio CT (DCT) es interesante, dado que es diferente en secuencia y estructura de acuerdo al tipo de sustrato que debe reconocer, por ejemplo, el DCT de las ACCasa, PCCasa, MCCasa, GCCasa, GDCasa y uno de los dominios de las MTCasa son similares, ya que reconocen a los acil-CoA (Hall *et al.* 2003; Zhang *et al.* 2003; Diacovich *et al.* 2004; Aguilar *et al.* 2006; Aguilar *et al.* 2008), este dominio se clasificó como CT tipo 1 (CT1). El DCT de las PCasa, ODCasa y el segundo dominio de las MTCasas son parecidos, y reconocen piruvato (Jitrapakdee *et al.* 1999); a este dominio se nombró tipo 2 (CT2). Existe un tercer dominio tipo 3 (CT3), el cual está presente solo en las UCasas (Figura 1) (Kanamori *et al.* 2004).

El dominio CT1 lleva a cabo la carboxilación de los acil-CoA para carboxilasas y transcarboxilasas (a), o en el caso de las descarboxilasas, la descarboxilación del sustrato para la formación de carboxi-biotina (b).



El dominio CT1 está formado por dos subdominios bien definidos, codificado por un solo gene en la mayoría de las carboxilasas, excepto las ACCasas de procariontas,

donde cada subdominio está codificado en genes por separado (Li y Cronan 1992; Jitrapakdee *et al.* 2003). Los dos subdominios comparten similitud en secuencia y estructura, lo cual ha llevado a sugerir que los dos subdominios surgieron de una duplicación génica en tándem de un gen antecesor único (Zhang *et al.* 2003; Diacovich *et al.* 2004).

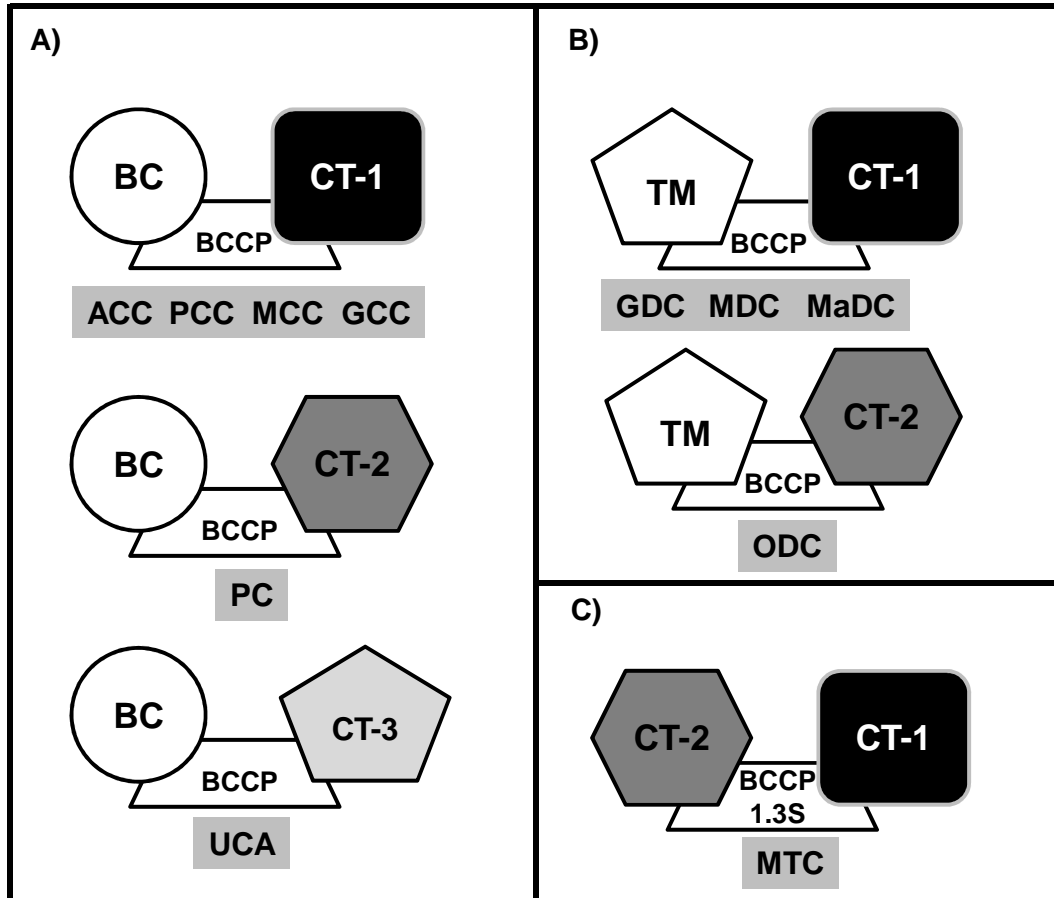


Figura 1.- Diagrama de los componentes de las tres clases de las CDB. A) La clase I agrupa a las enzimas carboxilasas. B) La clase II contiene a las enzimas descarboxilasas. C) La clase III está formada por las transcaboxilasas. Los diferentes dominios se esquematizan en diferente color y forma. BC-Dominio de carboxilación de biotina; CT-Dominio de carboxiltransferasa; BCCP-dominio acarreador de biotina; TM-dominio transmembranal.

A la fecha se conoce poco de la evolución del dominio CT1, sobre todo de la relación evolutiva que guardan los diferentes miembros de la familia de acuerdo a los acil-CoA que usan como sustrato. Por esto, este trabajo reporta el análisis filogenético del dominio CT1 de las CDB.

VIII.1.d) MATERIALES Y MÉTODOS

VIII.1.d.i) Obtención de las secuencias

Se utilizó una metodología similar a la reportada por Díaz-Pérez y col. (2007). Se realizó una búsqueda de homólogos mediante los programas Blastp y Psi-Blast (Altschul *et al.* 1997), utilizando la secuencia de aminoácidos de las subunidades β de la MCCasa (NP_250704) y de la GCCasa (NP_251578) de *P. aeruginosa*, las subunidades α (NP_752171) y β (AAA23801) de la ACCasa de *E. coli*. Se utilizó la base de datos nr (proteínas no redundantes del GeneBank), no se usó filtro de baja complejidad y se eliminaron las secuencias con un número E menor a 1×10^{-10} ; los parámetros de penalización de huecos (gaps) y la matriz usada fueron los de uso por omisión. Cuando se buscó con PSI-Blast, se utilizaron los mismos parámetros, repitiendo la búsqueda (iterando) hasta no encontrar más proteínas similares (convergencia), lo cual sucedió en la quinta iteración.

VIII.1.d.ii) Depuración de las secuencias

Se efectuó un alineamiento múltiple de todas las proteínas obtenidas en la búsqueda, utilizando el programa clustalW 2.0 (Larkin *et al.* 2007) con los parámetros por omisión. El alineamiento genera un árbol guía donde fácilmente se pueden ver las proteínas redundantes y los fragmentos; se utilizó TreeView (Page 1996), NJPlot (Perriere y Gouy 1996), y Dendroscope (Huson *et al.* 2007) para visualizar los árboles y hacer la selección de redundantes. Una vez que se seleccionaron las proteínas que posiblemente fueran redundantes y fragmentos, se procedió a hacer alineamientos por pares de cada grupo de proteínas sospechosas individualmente, para corroborar que efectivamente se tratara de proteínas repetidas y fragmentos. Estos pasos se repitieron cada vez que se eliminó alguna secuencia, hasta eliminar todas las proteínas redundantes y fragmentos. Se realizó un último paso de limpieza para verificar que no existieran fragmentos ni proteínas redundantes dentro del grupo final de trabajo.

VIII.1.d.iii) Análisis y corte de dominios

La localización y el análisis de dominios CT se realizó usando RPS-BLAST local. La base de datos de dominios CT conocidos fue obtenida a partir de la base de datos CDD (Conserved Domains Database) del NCBI (Marchler-Bauer *et al.* 2011). Los dominios conocidos usados fueron: COG4799 (Acetyl-CoA carboxylase, carboxyltransferase), COG0825 (AccA, Acetyl-CoA carboxylase alpha subunit), COG0777 (AccD, Acetyl-CoA carboxylase beta subunit), pfam01039 (Carboxyl_trans, Carboxyl transferase domain), pfam03255 (ACCA, Acetyl co-enzyme A carboxylase carboxyltransferase alpha subunit), pfam06833 (MdcE, Malonate decarboxylase gamma subunit), PRK05654 (acetyl-CoA carboxylase subunit beta), PRK12319 (acetyl-CoA carboxylase subunit alpha), PRK05724 (acetyl-CoA carboxylase carboxyltransferase), PRK07189 (malonate decarboxylase subunit beta) y CHL00174 (accD, acetyl-CoA carboxylase beta subunit). Los resultados de salida de RPS-Blast se usaron para ubicar los dominios CT en las secuencias del grupo de trabajo, y cortar los dominios para usarlos en el análisis filogenético. Para el análisis y corte de los dominios se usaron scripts escritos en Perl v5.0.

VIII.1.d.iv) Alineamiento múltiple

Para realizar el alineamiento del grupo final y de los subgrupos se utilizó el programa MAFFT (Kato y Toh 2008), el cual está diseñado para trabajar con grandes grupos de secuencias y cuenta con capacidad de uso de varios núcleos de procesamiento. Se usaron los parámetros FFT-NS-i y 100 iteraciones. El alineamiento obtenido se editó manualmente usando como editor de secuencias el programa BioEdit (Hall 1999).

VIII.1.d.v) Reducción del grupo de trabajo

Se efectuaron análisis de Blast-local de cada una de las secuencias del grupo de trabajo contra todo el grupo de trabajo. El resultado se analizó quitando las secuencias que tuvieran el porcentaje de identidad elegido, manteniendo la secuencia utilizada para la búsqueda.

VIII.1.d.vi) Construcción de los árboles filogenéticos

Para la construcción de las filogenias se utilizó el programa MEGA4 (Tamura *et al.* 2007). Se siguieron tres algoritmos distintos; de los métodos basados en matriz de peso (MMP) se usaron Vecino Más Cercano (NJ-Neighbor-Joining) y Mínima Evolución (ME) y de los métodos basados en características (MBC) se usó Máxima Parsimonia (MP). En la construcción de los árboles, se utilizó la delección pareada y como modelos evolutivos, la corrección de Poisson y la matriz JTT. Como prueba estadística se efectuó un análisis de Bootstrap de 500 a 1000 repeticiones para cada árbol. Se verificó la taxonomía del organismo de donde provenía cada secuencia para analizar los datos filogenéticos. Para la visualización y edición de los árboles se usaron los visualizadores de MEGA4 y de Dendroscope.

VIII.1.e) RESULTADOS Y DISCUSIÓN

VIII.1.e.i) Proteínas de la familia de transcarboxilasas dependientes de CoA (TDC)

El primer paso a seguir para generar la filogenia de una proteína es obtener un juego de datos de trabajo, que en este caso consistió en secuencias de proteínas homólogas. Para esto se realizó una búsqueda por similitud, obteniéndose 6111 secuencias (Tabla 2). Dado que existe una gran redundancia de información en las bases de datos, se realizó una limpieza en el grupo de secuencias, obteniendo un grupo final de 2556 secuencias no redundantes (Tabla 2).

Dentro del grupo, se encontraron proteínas de los tres dominios de la vida, archaeas (36 secuencias), bacterias (2074 secuencias) y eucariotes (446 secuencias). Se encontraron miembros distribuidos en dos reinos de archaeas (*Crenarchaeota* y *Euryarcheota*), en 25 reinos bacterianos y 11 reinos eucarióticos. La amplia distribución de los miembros de la familia nos indica que se trata de una familia ancestral.

Tabla 2.- Número de secuencias encontradas pertenecientes a la familia TDC

Grupo	Fuente	Número de secuencias
Búsqueda	PSI-BLAST (10)	4228
	BLAST (iterativo)	1883
Grupo Inicial		6111
Grupo Final		2556

Dentro del grupo de trabajo, se encontraron proteínas bacterianas cuya funcionalidad ya ha sido reportada, como son la ACCasa de *E. coli* (Li *et al.* 1992), la PCCasa de *Streptomyces coelicolor* (Diacovich *et al.* 2004), la MDCasa de *Veillonella parvula* (Huder y Dimroth 1993), la GDCasa de *Acidaminococcus fermentans* (Wendt *et al.* 2003) y la MTCasa de *Propionibacterium shermanii* (Hall *et al.* 2003); sin embargo, no se encontraron miembros de las PCasas, ODCasas y UCas. Para comprobar esto, se realizó un alineamiento por pares de las proteínas modelos CT, MCCasa y GCCasa de *P. aeruginosa*, contra las proteínas con función conocida reportadas

dentro de la familia de CDB (Jitrapakdee *et al.* 2003), no encontrándose identidad de secuencia con las enzimas que no reconocen compuesto con CoA (Tabla 3), es decir, los dominios CT1, son una familia independiente de los dominios CT2 y CT3. Además, se encontraron también miembros reportados como MaDCasas, las cuales descarboxilan al malonil, un substrato sin CoA, sin embargo, al verificar el mecanismo de reacción de estas enzimas, se encontró que las MaDCasas utilizan como grupo prostético un derivado de la CoA (2'-5-fosforribosil-3'-defosfo-CoA) para llevar a cabo su función (Dimroth *et al.* 1997). Por lo tanto, dado que todos los miembros obtenidos utilizan CoA o algún derivado, llamaremos a esta superfamilia transcarboxilasas dependientes de CoA (TDC).

Tabla 3.- Comparación de la subunidad CT (β) de la MCCasa y GCCasa de *P. aeruginosa* con otros miembros de la familia CDB.

Proteína	β -MCCasa	β -GCCasa
Acetil-CoA Carboxilasa	29	26
Propionil-CoA Carboxilasa	33	29
Metilcrotonil-CoA Carboxilasa	100	45
Geranil-CoA Carboxilasa	45	100
Metilmalonil-CoA Descarboxilasa	33	31
Glutaconil-CoA Descarboxilasa	24	23
Metilmalonil-CoA Transcarboxilasa	35	29
Piruvato Carboxilasa	NA	NA
Oxalacetato Descarboxilasa	NA	NA
Urea Carboxilasa	NA	NA

Valores en % de identidad

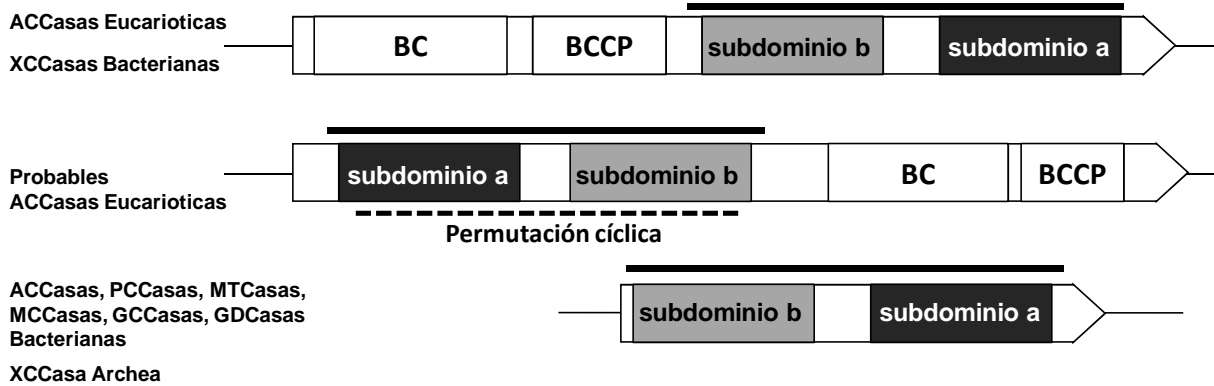
NA-No alinearon las secuencias

La mayoría de los dominios CT están formados por dos subdominios, b y a, cuyo ordenamiento génico se encuentra conservado (Figura 2). Se observó que un grupo de 30 probables ACCasas presentan un fenómeno de permutación cíclica del DCT, es decir, presentan un ordenamiento inverso de los subdominios que conforman al dominio (Figura 2), sin embargo, la funcionalidad de estas probables proteínas no ha sido caracterizada a nivel bioquímico. La conservación del ordenamiento en la

mayoría de los dominios de la familia sugiere una ventaja evolutiva de esta propiedad (Forslund *et al.* 2011), posiblemente para retener su función específica, para hacer más eficiente el plegamiento de las proteínas, o promover la interacción eficiente de las mismas.

Dado que las proteínas del grupo de trabajo además del dominio CT1 tienen otros dominios (Figura 1-2), los cuales no son informativos para llevar a cabo la filogenia de la familia, se procedió a llevar a cabo la localización de los dominios CT y a cortar las proteínas para realizar la filogenia de la familia de las TDC.

A)



B)

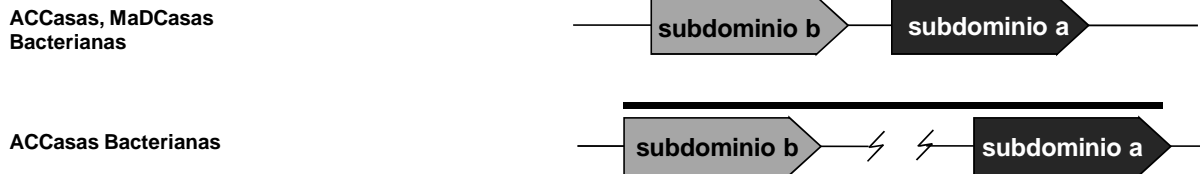


Figura 2.- Diagrama del ordenamiento de genes y dominios de las carboxilasas dependientes de biotina. El dominio CT está formado por dos subdominios, a y b. A) CDB que codifican el DCT en un solo gen. B) Carboxilasas dependientes de biotina que codifican los subdominios del dominio CT en genes independientes. Se marca el dominio CT con una línea continua y la permutación cíclica con línea punteada. Se marcan como XCCasas un grupo de CDB de bacterias multidominio cuya función es desconocida.

Al analizar la distribución de la longitud de las proteínas del grupo de trabajo (Figura 3), se observó que prácticamente todas las secuencias se agrupan en sólo dos regiones (250-450 aa y 475-750 aa), lo cual sugiere la presencia de sólo dos tipos de proteínas. El grupo de proteínas pequeñas, que corresponde a las proteínas que codifican a los subdominios a y b por separado (dominio separado –sCT), y el de las proteínas de cadena larga, que codifican al DCT en un solo gen (dominio completo –cCT). En ambos grupos existe una diferencia de tamaño promedio de 200 aa, la cual se debe a inserciones de secuencia en los extremos amino y carboxilo de las proteínas de mayor tamaño en cada grupo, además, en las cCT algunos miembros cuentan con una secuencia separadora entre los subdominios a y b de mayor tamaño. Para evitar la interferencia de las secuencias en los extremos de las proteínas de mayor tamaño, éstas fueron removidas antes de proceder a hacer los alineamientos finales.

VIII.1.e.ii) Análisis filogenético de los subdominios del DCT

Se encontró que, en general, las proteínas sCT se encuentran en parejas, con al menos un dominio a (CTa) y un dominio b (CTb) en los genomas de los organismos que las contienen; además, se encontraron 169 ejemplos de proteínas sCT que se encuentran adyacentes. Todas la CDB que se han caracterizado a la fecha requieren ambos subdominios para funcionar, a excepción de la MDCasa de *E. coli* YgfG (Haller *et al.* 2000), la cual es una proteína de 261 aa que no presenta similitud con las proteínas TDC. Esta observación sugiere que las proteínas TDC no son funcionales como monómeros. En las TDC que se han cristalizado, los subdominios b y a presentan la misma conformación tridimensional (plegamiento de crotonasa), y en algunas TDC aún se encuentra identidad de secuencia entre los dominio CTa y CTb, como la MCCasa de *P. aeruginosa* (NP_250705), lo que sugiere que las proteínas cCT surgieron de una fusión de las proteínas sCT.

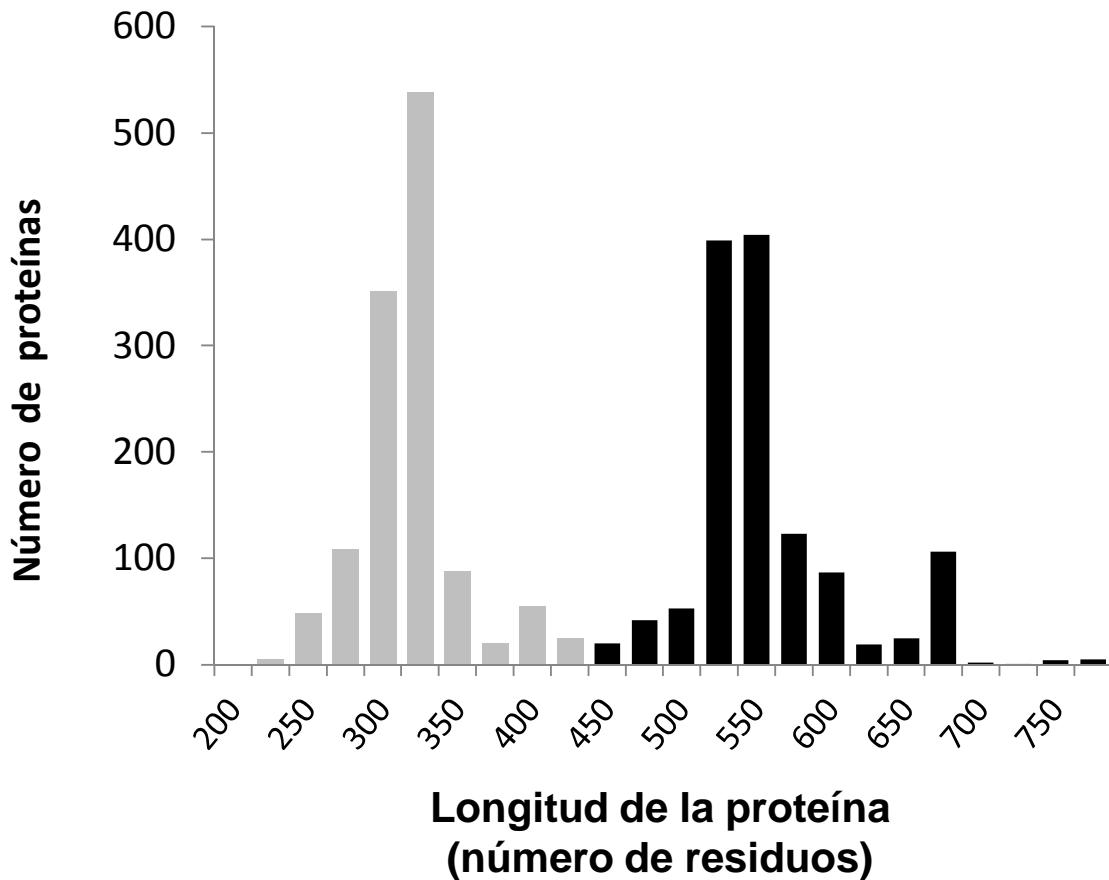


Figura 3.- Histograma de la longitud de las proteínas de la superfamilia de transcarboxilasas dependientes de CoA. Se observan dos grupos de proteínas, el primer grupo (barras grises) corresponde a las proteínas que codifican al dominio CT en genes independientes. El segundo grupo corresponde a las proteínas que codifican al dominio CT en un solo gen.

Para tratar de entender la relación entre las cCT y las sCT, se decidió hacer un análisis de los subdominios del DCT. Se llevó a cabo un análisis de la localización de los subdominios a y b dentro de las proteínas mayores, el cual se efectuó usando los resultados de RPS-Blast y se procedió a cortar los subdominios para tenerlos de manera individual, y se incluyeron con las sCT. Esto arrojó un grupo de trabajo de 3816 dominios, los cuales fueron alineados y se reconstruyó la filogenia de la familia.

Debido al gran número de secuencias, no fue posible realizar el análisis bootstrap con el grupo de dominios, obteniéndose únicamente un árbol general de este grupo

(Figura 4A). Para tener certeza de los resultados, se realizó un abordaje de disminución de secuencias, el cual consistió en hacer sub-juegos de secuencias con un 90% de similitud, y de cada grupo tomar una secuencia representativa. Las secuencias representativas se alinearon y se les realizó un análisis filogenético (Figura 4B). En el análisis se puede observar claramente que hay una separación de subdominios, lo que sugiere que ambos dominios han divergido, para formar familias distintas. Más aún, si la superfamilia de TDC surgió de un subdominio ancestral por un proceso de duplicación, este fenómeno no fue reciente.

Es interesante que en general las proteínas sCT se agrupen juntas (grupos I, II, VI y VII), separadas de las proteínas cCT (grupos III, IV, V, VIII y IX), y que existan solo ejemplos aislados de proteína cCT entre las sCT y viceversa. Estos datos sugieren un evento de divergencia ancestral entre las sCT y cCT.

Dada la distribución observada, se realizaron árboles de cada familia, CTa y CTb, de manera separada, para analizarlos a detalle. Los árboles obtenidos (Figura 5) son muy similares entre sí. La CTa familia cuenta con 5 grupos, nombrados del I al V, mientras la familia CTb cuenta con 4 grupos, nombrados del VI al IX. Los grupos presentan una distribución en espejo, es decir, cada grupo de la familia CTa tiene su contraparte en la familia CTb. La gran diferencia radica en la posición de un grupo de ACCasas eucarióticas, las cuales forman el grupo III en la familia CTa, mientras en la familia CTb forman parte del grupo IX (Figura 5). Esto no es sorprendente en las cCT, dado que al estar formando parte de un mismo gen, cuyos dominios interactúan, se esperarían cambios evolutivos similares (Yeang y Hausler 2007). Sin embargo, los genes que codifican para las proteínas sCT se encuentran juntas en forma de operón, o separadas en los genomas de los organismos que los contienen, y los grupos que contienen a cada dominio presentan la misma distribución, entonces, dado que las proteínas sCT que forman a cada subdominio interactúan para formar el DCT, esta presión de selección lleva a que las proteínas coevolucionen (Dandekar *et al.* 1998) y se encuentre la distribución observada. La distribución de los dominios en espejo sugiere que los subdominios de las TDC están sometidas a un proceso de

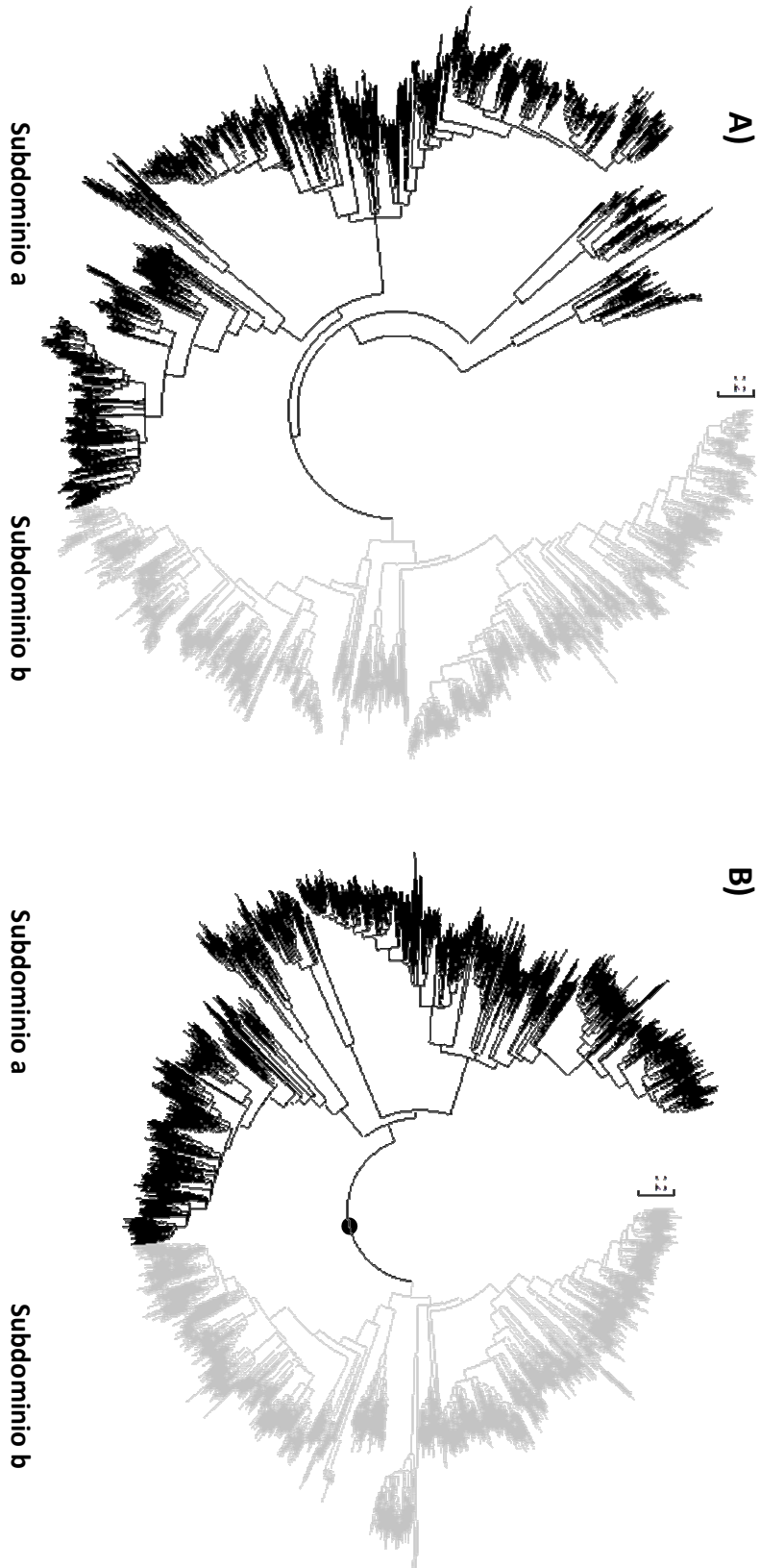


Figura 4.- Topología del árbol filogenético de los subdominios a y b de las superfamilia de TDC. A) Árbol filogenético del grupo de 3816 secuencias. Se realizó con el método de ME. B) Árbol filogenético del grupo reducido de secuencias. Se realizó con el método de ME y 500 repeticiones de bootstrap. En negro se señalan los grupos donde hay mayoritariamente subdominios a. En gris se marca los grupos donde hay mayoritariamente subdominios b. El círculo negro señala la separación de subdominios, con un bootstrap de 50%.

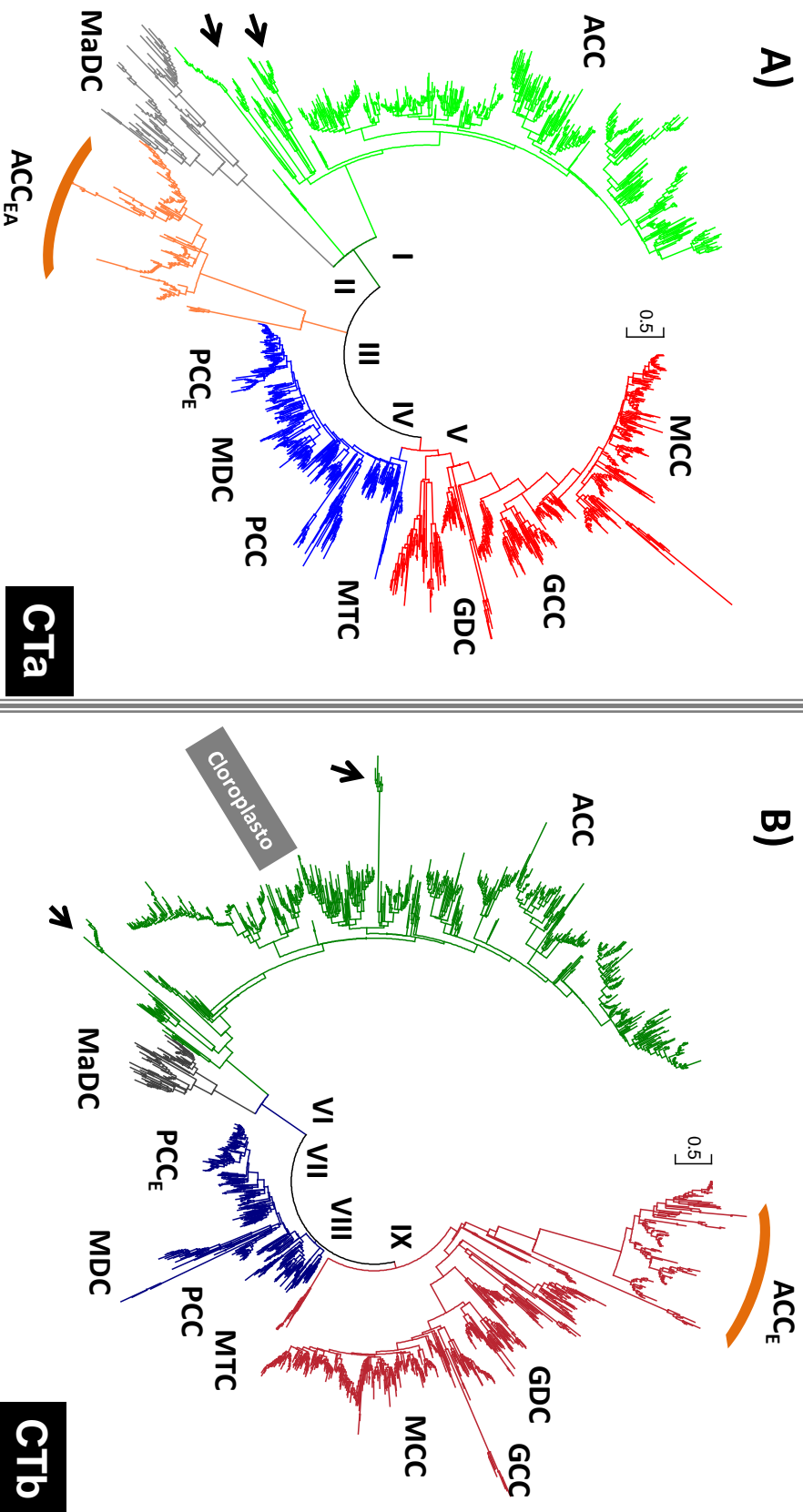


Figura 5 - Árbol filogenético de los subdominios a y b de las superfamilia de TDC. A) Árbol filogenético de la familia CTa. B) Árbol filogenético de la familia CTb. Los árboles se realizaron con el método de ME y 1000 repeticiones de bootstrap. Los números marcan los grupos principales encontrados. Se marcan con señal amarilla el grupo particular de cCT-ACCasas de eucariotas, y con marca gris, el grupo de sCT-ACCasas cloroplásticas. Las flechas señalan la localización de las secuencias con permutación cíclica.

coevolución (Pazos y Valencia 2008), más aún, la coevolución se ha observado en dominios proteicos que pertenecen a la misma proteína, o al mismo complejo proteico, y que llevan a cabo la misma función (Kim *et al.* 2006; Yeang *et al.* 2007), lo cual ocurre en las TDC.

Se encontraron 169 pares de proteínas sCT adyacentes. Los casos más claros son los de las MaDCasas (grupos II, CTa, y VII, CTb), las cuales forman 69 pares adyacentes, en distribución b-a (Figura 5). Los otros casos de proteínas adyacentes se presentaron en 100 ACCasas bacterianas.

También se encontró que existen ejemplos donde hay un número dispar de sCT, como en la bacteria *Hydrogenivirga sp.*, donde existen dos proteínas sCTa (ZP_02176657 y ZP_02179362), y una sola sCTb (ZP_02178314). Existen varias consideraciones a tomar en este caso, primero, los dos genes sCTa se agrupan muy cerca y codifican proteínas de tamaño similar (313 y 316 aa), por lo que la presencia de un gen extra pudiera ser debida a que se trata de una duplicación reciente, o que se trate de un pseudogen reciente (Kuo y Ochman 2010). Sin embargo, existe la posibilidad de que este evento sea un fenómeno regulatorio, donde la expresión de cada tipo de sCTa sea en condiciones específicas.

Otro caso interesante dentro de las sCT, es que se encontró un grupo de proteínas sCTb de plantas codificadas en el genoma de cloroplasto, localizadas en el grupo VI y que no presentan su contraparte sCTa (Figura 5). Analizando el genoma de *Arabidopsis thaliana* se encontró que presenta dos tipos de ACCasas, uno del tipo cCT multidominio (AAG40564), y un juego del tipo sCT, donde el dominio b se codifica en el cloroplasto (NP_051068) y el dominio a en el núcleo (NP_565880). La existencia del grupo mencionado, sin su par de dominio sCTa, puede deberse a que no se han secuenciado los genomas de esas plantas.

Enfocándonos en las cCT, se observa que el CTa se divide en tres grandes grupos, mientras el CTb se divide en dos. Las distancias evolutivas en los CTa son mayores

en relación a los tres grupos en comparación a los dos grupos CTb. Esto sugiere que, si bien ambos dominios coevolucionaron para conservar la interacción de los subdominios, la divergencia evolutiva entre ellos está presente, y en el caso de las ACCasa eucarióticas, es patente (Figura 5).

VIII.1.e.iii) Distribución funcional de las TDC

Existen varias proteínas con funcionalidad reportada en las CDB, las cuales se usaron para clasificar las posibles funciones dentro de los nueve grupos obtenidos (Tabla 4, Figura 5).

En general, se encontró una distribución de los grupos en relación a su funcionalidad (Tabla 4, Figura 5), salvo el caso de las ACCasas cCT ya mencionadas. Los grupos formados son homogéneos, las ACCasas bacterianas están separadas (grupo I y VI) de las ACCasas eucarióticas (grupos III y IX) y de las MaDCasas (grupo II y VII), las PCCasas, MDCasas y MTCasas forman otro grupo compacto (grupo IV y VIII), y las MCCasas, GCCasas y GDCasas forman otro grupo por separado (grupo V y IX). Esta distribución sugiere que las proteínas de la superfamilia TDC son parálogas entre sí, y es probable que hayan surgido de varios eventos de duplicación ancestrales.

En las TDC se han reportado casos de proteínas que son capaces de reconocer más de un sustrato, como es el caso de las GCCasa de *P. aeruginosa* (NP_251578) que reconoce geranil-CoA y metilcrotonil-CoA (grupo V y IX); la ACCasa (NP_626688) de *S. coelicolor*, la cual reconoce acetil-CoA, propionil-CoA y butiril-CoA (grupo I y VI); y la PCCasa de *S. coelicolor* (NP_629079) que reconoce propionil-CoA y butiril-CoA (grupo IV y VIII). Estas proteínas se encuentran en diferentes grupos, lo que sugiere que la promiscuidad, si bien se ha encontrado en pocos casos, está presente en toda la familia.

La distribución de las PCCasas, MDCasas y MTCasas en un grupo y de las MCCasas, GCCasas y GDCasas en otro grupo diferente, sugiere la especialización

Tabla 4.- Funciones, genes y organismo modelos de los nueve grupos de las TDC.

Grupo	Miembros ^a	Función reportada	Dominios ^b	Gen modelo	Organismo	Cita
I	594	Acetil-CoA Carboxilasa	B, E	NP_416819	<i>Escherichia coli</i>	Li et al. 1992
II	69	Malonil Descarboxilasa	B	AAC45403	<i>Malonomonas rubra</i>	Dimroth et al. 1997
III	143	Acetil-CoA carboxilasa	E	AAA20073	<i>Saccharomyces cerevisiae</i>	Al-Feel et al. 1992
IV	496	Metilmalonil-CoA Transcarboxilasa	A, B	Q8GBW6	<i>Propionibacterium freudenreichii</i>	Hall et al. 2003
		Propionil-CoA Carboxilasa	A, B, E	NP_629079	<i>Streptomyces coelicolor</i>	Diacovich et al. 2004
		Metilmalonil-CoA Descarboxilasa	A, B	AAC36820	<i>Veillonella parvula</i>	Huder y Dimroth 1993
V	540	Metilcrotonil-CoA Carboxilasa	B, E	NP_250754	<i>Pseudomonas aeruginosa</i>	Aguilar et al. 2008
		Geranil-CoA Carboxilasa	B, E	NP_251578	<i>Pseudomonas aeruginosa</i>	Aguilar et al. 2008
		Glutaconil-CoA Descarboxilasa	A, B	Q06700	<i>Acidaminococcus fermentans</i>	Wendt et al. 2003
VI	725	Acetil-CoA Carboxilasa	B, E	NP_416819	<i>Escherichia coli</i>	Li et al. 1992
VII	69	Malonil Descarboxilasa	B	AAC45402	<i>Malonomonas rubra</i>	Dimroth et al. 1997
VIII	461	Metilmalonil-CoA Transcarboxilasa	A, B	Q8GBW6	<i>Propionibacterium freudenreichii</i>	Hall et al. 2003
		Propionil-CoA Carboxilasa	A, B, E	NP_629079	<i>Streptomyces coelicolor</i>	Diacovich et al. 2004
		Metilmalonil-CoA Descarboxilasa	A, B	AAC36820	<i>Veillonella parvula</i>	Huder et al. 1993
IX	719	Metilcrotonil-CoA Carboxilasa	B, E	NP_250754	<i>Pseudomonas aeruginosa</i>	Aguilar et al. 2008
		Geranil-CoA Carboxilasa	B, E	NP_251578	<i>Pseudomonas aeruginosa</i>	Aguilar et al. 2008
		Glutaconil-CoA Descarboxilasa	A, B	Q06700	<i>Acidaminococcus fermentans</i>	Wendt et al. 2003
		Acetil-CoA carboxilasa	E	AAA20073	<i>Saccharomyces cerevisiae</i>	Al-Feel et al. 1992

^a Número de proteínas por grupo^b A-Dominio Archaea; B-Dominio Bacteria; E-Dominio Eukarya

de cada grupo proteico hacia los tipos de sustratos específicos, ya que las enzimas del grupo de las PCCasas carboxilan el carbono α del sustrato, mientras las enzimas del grupo de las MCCasas usan sustratos insaturados y llevan a cabo su función en el carbono γ .



VIII.1.e.iv) Transferencia de genes en la superfamilia

Se encontraron algunos ejemplos claros de transferencia horizontal de bacteria a eucariota. Dentro de las secuencias bacterianas sCT se localizaron los dominios de las proteínas eucarióticas de tipo cCT AAG35712 del protozooario *Giardia intestinalis* y XP_651829 de la ameba *Entamoeba histolytica* en el grupo I los dominios a y en el grupo VI los dominios b. Los dominios de estas proteínas eucarióticas se encuentran en grupos separados, lo que indica que no se trata de un fenómeno de pérdida de genes (Salzberg *et al.* 2001). Se analizó la secuencia de ambas proteínas para descartar que se tratara de contaminación bacteriana, fenómeno que ha ocurrido durante la secuenciación de algunos genomas eucarióticos (Gomez *et al.* 2005), el cual arrojó que es una proteína multidominio, parecida en ordenamiento a otras proteínas de eucariotes.

Un ejemplo de transferencia de genes conocido en eucariotes es la transferencia de genes del genoma cloroplástico al genoma nuclear en plantas (Martin *et al.* 2002; Richardson y Palmer 2007), como el caso de *A. thaliana* ya mencionado. Un caso similar ocurre en las TDC, en donde los genes sCT YP_002048770 (grupo VI) y YP_002049515 (grupo I) de la ameba fotosintética *Paulinella chromatophora* se agruparon con genes de cianobacterias. *P. chromatophora*, es un modelo de la transferencia endosimbiótica de genes (TEG) que provienen de cianobacterias (Nowack *et al.* 2011), por lo que podemos concluir que los genes mencionados tienen su origen en el endosimbionte bacteriano.

VIII.1.e.v) Origen evolutivo de las TDC

Las TDC forman una superfamilia ancestral que se encuentra ampliamente distribuida en los organismos, presenta distintos tipos de arreglos génicos y son codificadas en uno o más genes, por lo que su historia evolutiva es compleja.

Es posible encontrar similitud de secuencia en algunos subdominios CTa y CTb (ACCasa de la bacteria *Moraxella osloensis*, ABD36908; PCCasa de la bacteria *Corynebacterium accolens*, ZP_03932865; XCCasa de la bacteria *Corynebacterium*

jeikeium, YP_251457; MCCasa de la bacteria *P. aeruginosa*), además, se ha demostrado que la estructura tridimensional de ambos dominios es la misma (Hall *et al.* 2003; Wendt *et al.* 2003; Zhang *et al.* 2003; Diacovich *et al.* 2004; Huang *et al.* 2010), lo que sugiere que ambos subdominios han surgido de un evento de duplicación a partir de un dominio sCT ancestral.

Los árboles filogenéticos indican que todas las proteínas sCT son de origen bacteriano. En arqueas y en eucariotes solo se encuentran proteínas cCT (a excepción de las sCT adquiridas por TEG), mientras que las proteínas cCT de bacterias se agrupan con las cCT de eucariotes y arqueas, lo que sugiere que existen dos linajes claramente diferenciados en las TDC, el linaje sCT y el linaje cCT (Figura 6).

Una vez establecidos estos linajes primigenios, la distribución de los grupos parece indicar que las TDC se fueron especializando, adaptándose hacia nuevos sustratos. Los pocos ejemplos de promiscuidad observada en la superfamilia refuerza esta idea, ya que solo se ha visto promiscuidad hacia un grupo de sustratos de naturaleza similar, es decir la PCCasa de *S. coelicolor* usa sustratos libres en la posición α , y la GCCasa de *P. aeruginosa* usa sustratos ramificados, insaturados y ataca la posición γ (Figura 6).

VIII.1.f) CONCLUSIONES

Mediante el análisis bioinformático de las proteínas que forman el dominio CT se encontró que forman una superfamilia ancestral con al menos 2556 miembros reportados en los tres dominios de la vida. Todos los miembros de la familia son dependientes de CoA por lo que en este trabajo se nombra la superfamilia de transcarboxilasas dependientes de CoA (TDC). La superfamilia surgió a partir de una duplicación de dominios, seguida de un fenómeno de fusión, por lo que se formaron dos linajes ancestrales: el sCT, que agrupa a las proteínas no fusionadas, y el cCT que agrupa las proteínas fusionadas. Los posibles eventos evolutivos posteriores llevaron a formar familias de proteínas parálogas que se especializaron en tipos específicos de sustratos.

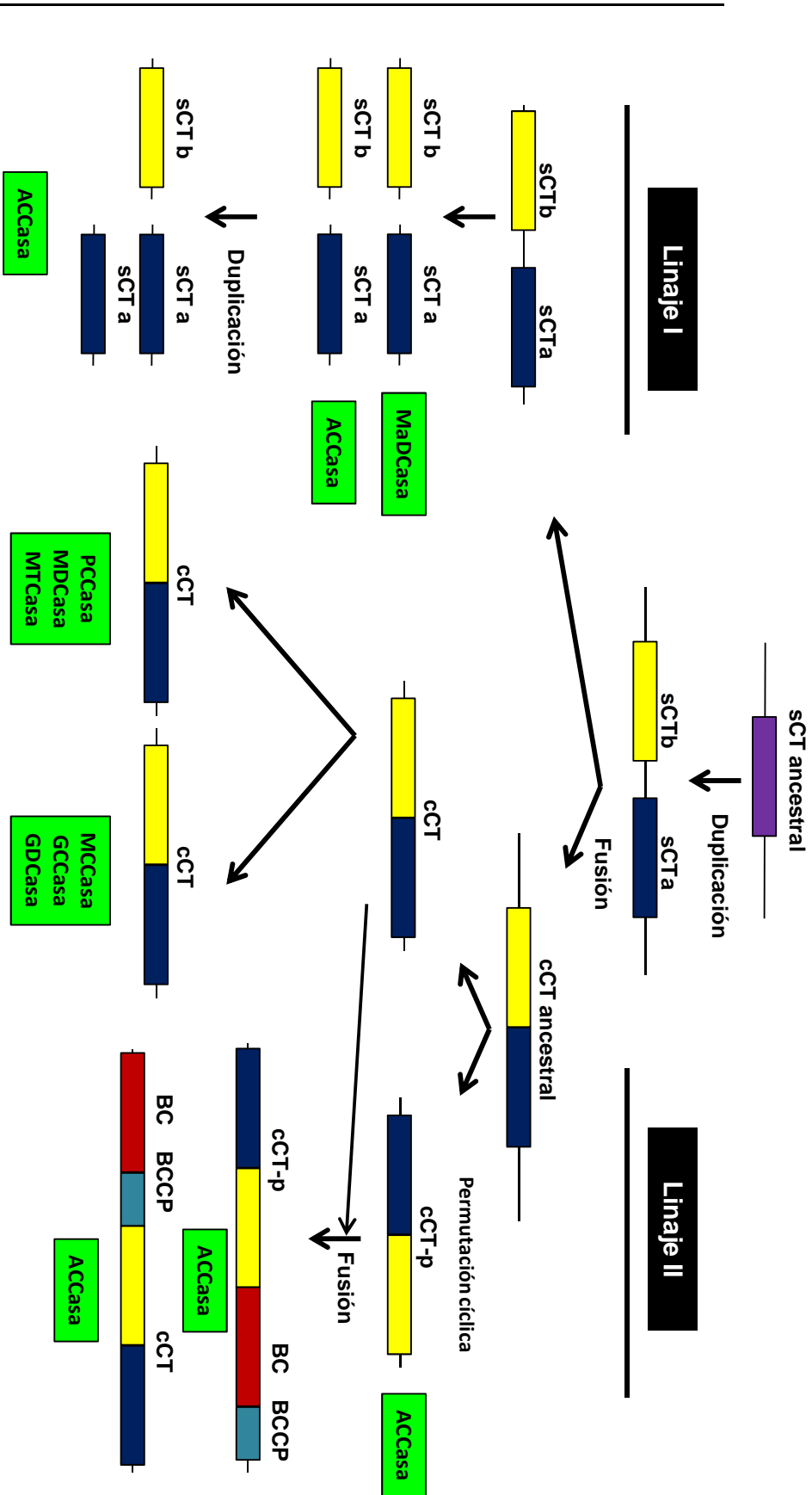


Figura 6.- Modelo evolutivo de la funcionalidad en la superfamilia TDC. Los dominios CTa, CTb, BC y BCCP se marcan en distintos colores. En recuadros se marcan las funciones de las proteínas de la superfamilia.

VIII.1.g) REFERENCIAS

- Aguilar, J. A., C. Díaz-Pérez, A. L. Díaz-Pérez, J. S. Rodríguez-Zavala, B. J. Nikolau y J. Campos-García. 2008. **Substrate specificity of the 3-methylcrotonyl coenzyme A (CoA) and geranyl-CoA carboxylases from *Pseudomonas aeruginosa***. J Bacteriol. 190(14): 4888-4893.
- Aguilar, J. A., A. N. Zavala, C. Díaz-Pérez, C. Cervantes, A. L. Díaz-Pérez y J. Campos-García. 2006. **The *atu* and *liu* clusters are involved in the catabolic pathways for acyclic monoterpenes and leucine in *Pseudomonas aeruginosa***. Appl Environ Microbiol. 72(3): 2070-2079.
- Al-Feel, W., S. S. Chirala y S. J. Wakil. 1992. **Cloning of the yeast FAS3 gene and primary structure of yeast acetyl-CoA carboxylase**. Proc Natl Acad Sci U S A. 89(10): 4534-4538.
- Altschul, S. F., T. L. Madden, A. A. Schaffer, J. Zhang, Z. Zhang, W. Miller y D. J. Lipman. 1997. **Gapped BLAST and PSI-BLAST: a new generation of protein database search programs**. Nucleic Acids Res. 25(17): 3389-3402.
- Buckel, W. 2001. **Sodium ion-translocating decarboxylases**. Biochim Biophys Acta. 1505(1): 15-27.
- Buckel, W. 2001. **Unusual enzymes involved in five pathways of glutamate fermentation**. Appl Microbiol Biotechnol. 57(3): 263-273.
- Dandekar, T., B. Snel, M. Huynen y P. Bork. 1998. **Conservation of gene order: a fingerprint of proteins that physically interact**. Trends Biochem Sci. 23(9): 324-328.
- Diacovich, L., D. L. Mitchell, H. Pham, G. Gago, M. M. Melgar, C. Khosla, H. Gramajo y S. C. Tsai. 2004. **Crystal structure of the beta-subunit of acyl-CoA carboxylase: structure-based engineering of substrate specificity**. Biochemistry. 43(44): 14027-14036.
- Díaz-Pérez, C., C. Cervantes, J. Campos-García, A. Julian-Sánchez y H. Riveros-Rosas. 2007. **Phylogenetic analysis of the chromate ion transporter (CHR) superfamily**. FEBS J. 274(23): 6215-6227.

- Dimroth, P. y H. Hilbi. 1997. **Enzymic and genetic basis for bacterial growth on malonate.** Mol Microbiol. 25(1): 3-10.
- Dimroth, P., P. Jockel y M. Schmid. 2001. **Coupling mechanism of the oxaloacetate decarboxylase Na(+) pump.** Biochim Biophys Acta. 1505(1): 1-14.
- Forslund, K., I. Pekkari y E. L. Sonnhammer. 2011. **Domain architecture conservation in orthologs.** BMC Bioinformatics. 12: 326.
- Gomez, S. M., K. Eiglmeier, B. Segurens, P. Dehoux, A. Couloux, C. Scarpelli, P. Wincker, J. Weissenbach, P. T. Brey y C. W. Roth. 2005. **Pilot Anopheles gambiae full-length cDNA study: sequencing and initial characterization of 35,575 clones.** Genome Biol. 6(4): R39.
- Hall, P. R., Y. F. Wang, R. E. Rivera-Hainaj, X. Zheng, M. Pustai-Carey, P. R. Carey y V. C. Yee. 2003. **Transcarboxylase 12S crystal structure: hexamer assembly and substrate binding to a multienzyme core.** EMBO J. 22(10): 2334-2347.
- Hall, T. A. 1999. **BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT.** Nucleic Acids Symp Ser 41: 95-98.
- Haller, T., T. Buckel, J. Retej y J. A. Gerlt. 2000. **Discovering new enzymes and metabolic pathways: conversion of succinate to propionate by Escherichia coli.** Biochemistry. 39(16): 4622-4629.
- Houwen, F. P., J. Plokker, A. J. M. Stams y A. J. B. Zehnder. 1990. **Enzymatic evidence for involvement of the methylmalonyl-CoA pathway in propionate oxidation by Syntrophobacter wolinii.** Archives of Microbiology. 155(1): 52-55.
- Huang, C. S., K. Sadre-Bazzaz, Y. Shen, B. Deng, Z. H. Zhou y L. Tong. 2010. **Crystal structure of the alpha(6)beta(6) holoenzyme of propionyl-coenzyme A carboxylase.** Nature. 466(7309): 1001-1005.
- Huder, J. B. y P. Dimroth. 1993. **Sequence of the sodium ion pump methylmalonyl-CoA decarboxylase from Veillonella parvula.** J Biol Chem. 268(33): 24564-24571.



- Huson, D. H., D. C. Richter, C. Rausch, T. DeZulian, M. Franz y R. Rupp. 2007. **Dendroscope: An interactive viewer for large phylogenetic trees.** BMC Bioinformatics. 8: 460.
- Jitrapakdee, S. y J. C. Wallace. 1999. **Structure, function and regulation of pyruvate carboxylase.** Biochem J. 340 (Pt 1): 1-16.
- Jitrapakdee, S. y J. C. Wallace. 2003. **The biotin enzyme family: conserved structural motifs and domain rearrangements.** Curr Protein Pept Sci. 4(3): 217-229.
- Kanamori, T., N. Kanou, H. Atomi y T. Imanaka. 2004. **Enzymatic characterization of a prokaryotic urea carboxylase.** J Bacteriol. 186(9): 2532-2539.
- Katoh, K. y H. Toh. 2008. **Recent developments in the MAFFT multiple sequence alignment program.** Brief Bioinform. 9(4): 286-298.
- Kim, Y., M. Koyuturk, U. Topkara, A. Grama y S. Subramaniam. 2006. **Inferring functional information from domain co-evolution.** Bioinformatics. 22(1): 40-49.
- Kuo, C. H. y H. Ochman. 2010. **The extinction dynamics of bacterial pseudogenes.** PLoS Genet. 6(8).
- Larkin, M. A., G. Blackshields, N. P. Brown, R. Chenna, P. A. McGettigan, H. McWilliam, F. Valentin, I. M. Wallace, A. Wilm, R. Lopez, J. D. Thompson, T. J. Gibson y D. G. Higgins. 2007. **Clustal W and Clustal X version 2.0.** Bioinformatics. 23(21): 2947-2948.
- Li, S. J. y J. E. Cronan, Jr. 1992. **The genes encoding the two carboxyltransferase subunits of Escherichia coli acetyl-CoA carboxylase.** J Biol Chem. 267(24): 16841-16847.
- Marchler-Bauer, A., S. Lu, J. B. Anderson, F. Chitsaz, M. K. Derbyshire, C. DeWeese-Scott, J. H. Fong, L. Y. Geer, R. C. Geer, N. R. Gonzales, M. Gwadz, D. I. Hurwitz, J. D. Jackson, Z. Ke, C. J. Lanczycki, F. Lu, G. H. Marchler, M. Mullokandov, M. V. Omelchenko, C. L. Robertson, J. S. Song, N. Thanki, R. A. Yamashita, D. Zhang, N. Zhang, C. Zheng y S. H. Bryant. 2011.

CDD: a Conserved Domain Database for the functional annotation of proteins. Nucleic Acids Res. 39(Database issue): D225-229.

Martin, W., T. Rujan, E. Richly, A. Hansen, S. Cornelsen, T. Lins, D. Leister, B. Stoebe, M. Hasegawa y D. Penny. 2002. **Evolutionary analysis of Arabidopsis, cyanobacterial, and chloroplast genomes reveals plastid phylogeny and thousands of cyanobacterial genes in the nucleus.** Proc Natl Acad Sci U S A. 99(19): 12246-12251.

Nowack, E. C., H. Vogel, M. Groth, A. R. Grossman, M. Melkonian y G. Glockner. 2011. **Endosymbiotic gene transfer and transcriptional regulation of transferred genes in Paulinella chromatophora.** Mol Biol Evol. 28(1): 407-422.

Page, R. D. 1996. **TreeView: an application to display phylogenetic trees on personal computers.** Comput Appl Biosci. 12(4): 357-358.

Pazos, F. y A. Valencia. 2008. **Protein co-evolution, co-adaptation and interactions.** EMBO J. 27(20): 2648-2655.

Perriere, G. y M. Gouy. 1996. **WWW-query: an on-line retrieval system for biological sequence banks.** Biochimie. 78(5): 364-369.

Richardson, A. O. y J. D. Palmer. 2007. **Horizontal gene transfer in plants.** J Exp Bot. 58(1): 1-9.

Salzberg, S. L., O. White, J. Peterson y J. A. Eisen. 2001. **Microbial genes in the human genome: lateral transfer or gene loss?** Science. 292(5523): 1903-1906.

Tamura, K., J. Dudley, M. Nei y S. Kumar. 2007. **MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0.** Mol Biol Evol. 24(8): 1596-1599.

Thompson, J. D., D. G. Higgins y T. J. Gibson. 1994. **CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice.** Nucleic Acids Res. 22(22): 4673-4680.

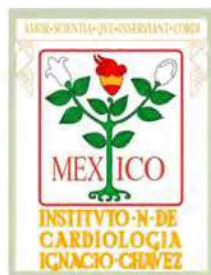
- Wendt, K. S., I. Schall, R. Huber, W. Buckel y U. Jacob. 2003. **Crystal structure of the carboxyltransferase subunit of the bacterial sodium ion pump glutaconyl-coenzyme A decarboxylase.** EMBO J. 22(14): 3493-3502.
- Westover, J. B., S. I. Goodman y F. E. Frerman. 2001. **Binding, hydration, and decarboxylation of the reaction intermediate glutaconyl-coenzyme A by human glutaryl-CoA dehydrogenase.** Biochemistry. 40(46): 14106-14114.
- Yeang, C. H. y D. Haussler. 2007. **Detecting coevolution in and among protein domains.** PLoS Comput Biol. 3(11): e211.
- Zhang, H., Z. Yang, Y. Shen y L. Tong. 2003. **Crystal structure of the carboxyltransferase domain of acetyl-coenzyme A carboxylase.** Science. 299(5615): 2064-2067.

World Journal of Microbiology and Biotechnology

ISSN: 0959-3993
2012
Volume 28
Number 3
pp 1185-1191



DOI: 10.1007/s11274-011-0921-1



César Díaz-Pérez
José Salud Rodríguez-Zavala
Alma Laura Díaz-Pérez
Jesús Campos-García

Capítulo II:

"Co-expression of α and β subunits of the 3-Methyl-crotonyl-Coenzyme A carboxylase from *Pseudomonas aeruginosa*"

VIII.2) Capítulo II:

Co-expression of α and β subunits of the 3-Methyl-crotonyl-Coenzyme A carboxylase from *Pseudomonas aeruginosa*

World Journal of Microbiology and Biotechnology., (2012) 28:1185–1191.

doi:10.1007/s11274-011-0921-1.

En *P. aeruginosa*, la enzima 3-metilcrotonil-CoA carboxilasa (MCCasa, E.C.6.4.1.4) es el punto de convergencia de las rutas de degradación de los monoterpenos acíclicos de la familia del citronelol y de la asimilación de leucina. Mutantes insercionales en los genes que codifican a la MCCasa son incapaces de crecer en leucina o citronelol como única fuente de carbono, demostrando la importancia de la enzima (Aguilar *et al.*, 2006; Díaz-Pérez *et al.*, 2004). La MCCasa pertenece a la familia de carboxilasas dependientes de biotina (CDB); es una proteína formada por dos subunidades codificados por los genes *liuD* (subunidad α) y *liuB* (subunidad β). Dado que las biotin carboxilasas catalizan reacciones claves en el metabolismo (Chapman-Smith *et al.*, 2001), la MCCasa ha sido considerada como un punto importante en ambas vías catabólicas.

En nuestro grupo de trabajo se comenzó la caracterización bioquímica de ambas rutas metabólicas. Para caracterizar cinéticamente a la MCCasa, se procedió a clonar los dos genes que la codifican en plásmidos separados, para sobre-expresar y purificar cada una de las subunidades. Posteriormente se desnaturalizaron simultáneamente ambas subunidades y se renaturalizaron conjuntamente para la reconstitución y formación de la estructura cuaternaria activa de la enzima (Aguilar *et al.*, 2008). Esta metodología es laboriosa, por lo que se consideró necesario buscar alternativas eficaces que permitan obtener a la MCCasa activa para su caracterización bioquímica.

En este trabajo se realizó la co-expresión de las dos subunidades de la MCCasa, a partir de un solo plásmido. En este proceso se realizó una fusión traduccional de una etiqueta de seis histidinas al gen que codifica la subunidad α de la MCCasa, por lo

que fue posible co-purificar a la enzima en un sólo paso, mediante cromatografía de afinidad, con un grado de pureza mayor al 95%. La caracterización enzimática arrojó que la enzima tiene un pH óptimo de 9.0 y una temperatura óptima de 30°C. La MCCasa co-expresada mostró un comportamiento michaeliano, es decir, presentó una cinética hiperbólica, con una Km y una Vmax de 168 μM para el MC-CoA y 430 $\text{nmoles}\cdot\text{mg}^{-1}\cdot\text{min}^{-1}$. Se demostró que la enzima forma un heterómero activo $\alpha_4\beta_4$. La enzima es activa durante 12 horas. Se encontró que conforme pasa el tiempo de almacenamiento, la MCCasa pierde su estado de oligomerización y su actividad.

World J Microbiol Biotechnol (2012) 28:1185–1191
 DOI 10.1007/s11274-011-0921-1

ORIGINAL PAPER

Co-expression of α and β subunits of the 3-methylcrotonyl-coenzyme A carboxylase from *Pseudomonas aeruginosa*

César Díaz-Pérez · José Salud Rodríguez-Zavala ·
 Alma Laura Díaz-Pérez · Jesús Campos-García

Received: 1 August 2011 / Accepted: 11 October 2011 / Published online: 20 October 2011
 © Springer Science+Business Media B.V. 2011

Abstract *Pseudomonas aeruginosa* is a versatile bacterium that can grow using citronellol or leucine as sole carbon source. For both compounds the degradation pathways converge at the key enzyme 3-methylcrotonyl coenzyme-A carboxylase (MCCase). This enzyme is a complex formed by two subunits (α and β), encoded by the *liuD* and *liuB* genes, respectively; both are essential for enzyme function. Previously, both subunits had been separately expressed and then the complex re-constituted, however this methodology is laborious and produces low yield of active enzyme. In this work, the MCCase subunits were co-expressed in the same plasmid and purified in one step by affinity chromatography using the LiuD-His tag protein, interacting with the LiuB-S tag recombinant protein. The purified enzyme lost most of the activity within few hours of storage. The co-expressed subunits formed an $(\alpha\beta)_4$ complex that suffered a modification of its oligomerization state after storage, which probably contributed to the loss on activity observed. The recombinant MCCase enzyme presented optimum pH and temperature values of 9.0 and 30° C, respectively. Functionally, MCCase showed Michaelian kinetics behavior with a K_m for its substrate and V_{max} of 168 μ M and 430 nmoles $mg^{-1}min^{-1}$,

respectively. The results suggest that the co-expression and co-purification of the subunits is a suitable procedure to obtain the active complex of the MCCase from *Pseudomonas aeruginosa* in a single step.

Keywords *Pseudomonas aeruginosa* · 3-methylcrotonyl-CoA carboxylase · Citronellol · Leucine · Enzyme reconstitution


Introduction

In *Pseudomonas aeruginosa*, the acyclic monoterpenes of the citronello family (AMTC) and leucine/isovalerate catabolic pathways converge in the metabolite 3-methylcrotonyl-CoA (MC-CoA) (Fig. 1; Campos-García 2010). The enzyme that uses the MC-CoA to produce 3-methylglutaconyl-CoA is the 3-methylcrotonyl-CoA carboxylase (MCCase, EC 6.4.1.4) (Aguilar et al. 2006), a member of the biotin-dependent enzymes family (Jitrapakdee and Wallace 2003). The MCCase is a key enzyme in leucine and AMTC catabolism when they are used as sole carbon source (Díaz-Pérez et al. 2004; Aguilar et al. 2006). The MCCase is composed by the α - and β -subunits, which are encoded by the *liuD* (PA2012) and *liuB* (PA2014) genes, respectively. Both subunits are essential for enzyme function (Hoschle et al. 2005; Aguilar et al. 2006). The α -subunit is constituted by the biotin-carboxylase domain (BC) and the biotin-carrier domain (BCCP); biotin carboxylation takes place in this subunit. The β -subunit contains the transcarboxylase domain, involved in the transference of the carboxyl-group to the MC-CoA substrate. It has been described that homologues of this enzyme occur in many species, forming a hetero-octameric oligomer $(\alpha\beta)_4$, like the MCCases from *P. citronellolis*,

Electronic supplementary material The online version of this article (doi:10.1007/s11274-011-0921-1) contains supplementary material, which is available to authorized users.

C. Díaz-Pérez · A. L. Díaz-Pérez · J. Campos-García (✉)
 Instituto de Investigaciones Químico Biológicas, Universidad Michoacana de San Nicolás de Hidalgo, Edificio B-3, Ciudad Universitaria, 58030 Morelia, Michoacán, Mexico
 e-mail: jcgarcia@umich.mx

J. S. Rodríguez-Zavala
 Departamento de Bioquímica, Instituto Nacional de Cardiología, México, D.F., Mexico

 Springer



Achromobacter sp., *Pisum sativum*, and *Solanum tuberosum* (Schiele et al. 1975; Fall and Hector 1977; Alban et al. 1993; Wurtele and Nikolau 2000), whereas that from *Zea mays* forms a hetero-dodecamer ($\alpha\beta$)₆ (Diez et al. 1994).

To understand better the assembly of the active complex of the MCCase, our group described previously the functional characterization of the MCCase (Aguilar et al. 2008). The α and β subunits were separately expressed and purified; then, the subunits were mixed, denatured, and

re-natured to reconstitute the active MCCase enzyme complex (Aguilar et al. 2008). This strategy has the disadvantages of being laborious and to produce a small amount of active enzyme. One alternative that has been followed for the characterization of some cozymes, conformed by two subunits, involves their co-expression. It requires the simultaneous expression of both genes in a co-expression vector, which is then transformed into an *Escherichia coli* strain to obtain the functional enzyme from the cells-free extract. Further, the co-expression strategy may increase the enzyme yield because the complex can be purified without having to be reconstituted in one step (Kerrigan et al. 2011). This strategy has been successfully applied for the purification of the homologous biotin-containing enzyme geranyl-CoA carboxylase (GCCase) (Aguilar et al. 2008). The goal of this study was to implement a more efficient expression and purification procedure for the functional characterization of the MCCase enzyme complex. We describe herein the one step expression, the purification and biochemical characterization of the MCCase from *P. aeruginosa*.

Materials and methods

Bacterial strains, plasmids, and culture conditions

The bacterial strains used in this work were *Escherichia coli* JM101 and *E. coli* BL21 (DE3) pLysS. Plasmids used were pGEM-T Easy (Promega) and pCDFDuet-1 (Novagen). The strains were grown at 30°C in Luria-Bertani broth (LB) or 2xYT expression-medium (2xYT, yeast extract 10 g/l, tryptone 16 g/l, NaCl 5 g/l). Solid media were prepared by adding agar 15 g/l. Antibiotics used were streptomycin at 30 µg/ml and chloramphenicol at 30 µg/ml.

Cloning *liuD* and *liuB* genes in dual expression vector

General molecular genetic techniques were used according to standard protocols (Sambrook et al. 2001). *P. aeruginosa* DNA genomic was used as template for PCR amplification. For *liuD* gene amplification, the LiuD1 oligonucleotide (5'-AGGATCCCATGAGCGAATTCCAGACCATCCAG-3') was used to introduce a *Bam*HI restriction site upstream of the start codon, and LiuD2 (5'-GAAGCTTTTGGTCGTTTCTCCAGGCGGG-3') was used to introduce a *Hind*III restriction site at the 3' end of *liuD* gene. For *liuB* gene, amplification strategy was the same, using the oligonucleotides LiuB1 (5'-CAGATCTCATGCCATCCTGCACACCCAGATC-3'), introducing a *Bgl*II site and LiuB2 (5'-CGGTACCCATGCCGAACACGCCGAAGGC-3'), introducing a *Kpn*I; each

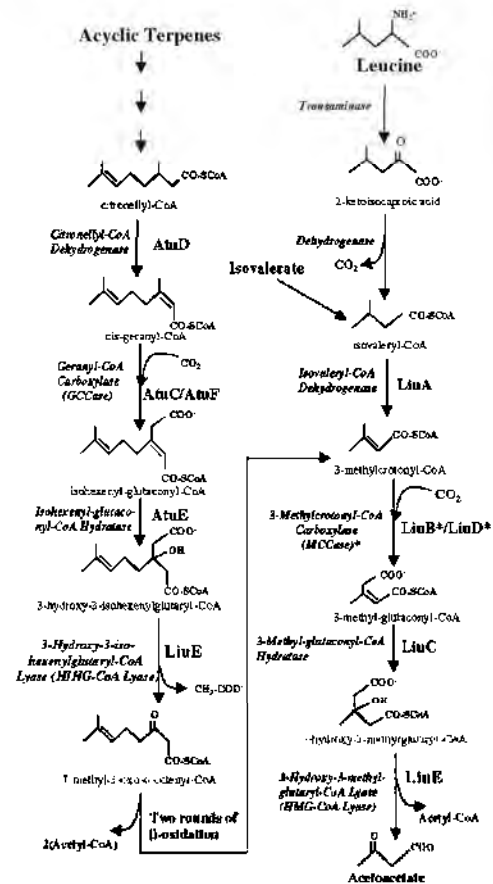


Fig. 1 Role of the 3-methylcrotonyl-CoA carboxylase from *P. aeruginosa* on the leucine/isovalerate and acyclic terpene pathways. AtuD, citronellyl-CoA dehydrogenase; AtuE, isohexenyl-glutaconyl-CoA hydratase; LiuA, isovaleryl-CoA dehydrogenase; LiuB/LiuD, MCCase; LiuC, 3-methyl-glutaconyl-CoA hydratase; LiuE, promiscuous 3-hydroxy-3-methyl-glutaryl-CoA and 3-hydroxy-3-isohexenyl-glutaryl-CoA lyase. Asterisks indicate the enzyme in study. (adapted from Aguilar et al. 2006; Aguilar et al. 2008; Chávez-Avilés et al. 2009; Campos-García 2010)

resulting PCR fragment was cloned into pGem-T Easy vector. Fragments were cut and subcloned into the pCDFDuet-1 co-expression vector, given the pCDB3 plasmid (Supplementary material). The pCDB3 plasmid was used for to transform electrocompetent *E. coli* JM101 cells, and subsequently analyzed by digestion with restriction endonucleases and confirming by DNA sequencing.

Expression and purification of recombinant MCCase enzyme

The pCDB3 plasmid containing the *liuD* and *liuB* genes was transferred into *E. coli* BL21 strain, and cultured in 400 ml of 2xYT medium for 4 h at 37°C with shaking (150 rpm). Enzyme expression was induced in the bacterial culture with 0.1 mM of isopropyl- β -D-thiogalactopyranoside (IPTG), and the cells were incubated for 12 h at 37°C with shaking. Cells were harvested by centrifugation to 5,000 \times g for 10 min at 4°C. The bacterial pellet was suspended in 10 ml of buffer A (50 mM Tris-HCl, 500 mM NaCl, pH 7.5) and disrupted by sonication at 4°C. Cell extract was centrifuged to 15,000 \times g for 30 min at 4°C to eliminate undisturbed cells and cell debris. Recombinant proteins were purified from crude extracts according to the TALON™ purification kit protocol (DB-Biosciences). The resin-protein mixture was washed four times with 1 ml of wash buffer (buffer A plus 20 mM imidazole), and after with 500 μ l of elution buffer (buffer A containing 250 mM imidazole) to elute the protein bound. The purified MCCase enzyme was diluted 1:4 (v/v) with buffer A plus 50% (v/v) glycerol. Protein concentration was determined by the Lowry method (Lowry et al. 1951).

Western-blot analysis

Purified proteins (10 mg) were separated by SDS-PAGE (10% polyacrylamide) and electrophoretically transferred to nitrocellulose membrane (Amersham Biosciences). MCCase α subunit was detected using the avidin-horseradish peroxidase (HRP) conjugate (Bio-Rad) as previously described (Aguilar et al. 2006). MCCase β subunit was detected using anti-S-Tag HRP conjugate or using polyclonal mouse anti-LiuB antibody. Polyclonal mouse anti-LiuB antibody was obtained from a polyclonal serum after immunization of mice with LiuB-His recombinant purified protein. The polyclonal serum was obtained at the Iowa State University Hybridoma Facility (www.biotech.iastate.edu/service_facilities/hybridoma.html). The membranes were probed using the polyclonal serum, and the anti-mouse-HRP conjugate was used as the second antibody (donkey anti-mouse immunoglobulin G-HRP; Santa Cruz Biotechnology,

Inc.) as indicated by the provider. HRP color development was carried out using 4-chloro-1-naphthol (Sigma) and hydrogen peroxide.

Molecular weight determination

The native molecular weight of the purified MCCase enzyme was determined by using an 80 cm-long column packed with Sephacryl-S-300. A 500 μ l sample of the MCCase purified fraction without glycerol was charged into the column, and eluted with buffer containing 100 mM Na₂HPO₄, 100 mM NaCl and 0.025% (v/v) β -mercaptoethanol, pH 7.4. The following proteins were used as molecular weight markers (MW): thyroglobulin, 660 kDa; ferritin, 440 kDa; catalase, 232 kDa and aldolase, 158 kDa.

Geranyl-CoA synthesis

Synthesis of geranyl-CoA (G-CoA) was performed by the mixed-anhydride method with slight variations (Guan et al. 1999). Geranic acid (770 μ mol) was dissolved in 5.1 ml tetrahydrofuran and neutralized by adding an equimolar amount of triethylamine. Ethylchloroformate (770 μ mol) was added, and the mixture was stirred for 30 min at room temperature and filtered. The filtrate containing the anhydride was added dropwise to 29 μ mol of coenzyme A, dissolved in 12 ml of water:tetrahydrofuran (3:2 v/v) solution and neutralized to pH 8.0 with solid NaHCO₃. The mixture was stirred for 25 min, and then 4 ml of water was added and the pH adjusted to 3.0 with 2 N HCl. The solution was extracted three times with diethyl ether. The aqueous phase was lyophilized. The lyophilized G-CoA was dissolved in MOPS 100 mM, pH 7.8, and its concentration was determined by the hydroxamate method (Lipmann and Tuttle 1945).

Carboxylase activity determination

MCCase and GCCase enzymatic activities for the purified proteins were measured by the incorporation of ¹⁴C into acid nonvolatile material as previously described (Hector and Fall 1976; Aguilar et al. 2008). The reaction mixture contained 100 mM K₂HPO₄ (pH 9.0), 100 mM Tris-HCl, 10 mM ATP, 10 mM MgCl₂, 50 mM NaH¹⁴CO₃ (specific activity, 1.96 GBq/mmol; 53 mCi/mmol; Amersham), 10 μ g of protein, and variable concentrations of the MC-CoA (Sigma) or G-CoA, in 100 μ l total reaction volume. The reaction mix was incubated 10 min without NaH¹⁴CO₃. The reaction was started by the addition of NaH¹⁴CO₃, the mixture was incubated at 30°C; after 10 min the reaction was stopped by adding 200 μ l of 6 M HCl. The reaction mix was evaporated to dryness at 94°C,



and the residue was suspended in 100 μ l of distilled water. Radioactivity was quantified using a liquid scintillation counter (Beckman LS Analyzer LS6000IC). Non-specific ^{14}C CO₂ fixation by purified protein was assayed in the absence of MC-CoA or G-CoA. For the pH optimum assay a buffer mixture with 20 mM NaHPO₄, 20 mM Tris, 20 mM diethanolamine and 20 mM diethiamide was used and pH was adjusted to 7.0, 7.5, 8.0, 8.5, and 9.0. For optimal temperature assay, water baths at temperatures of 20°, 25°, 30°, 35° and 40°C were used.

Results and discussion

Construction of the co-expression plasmid containing the α and β subunits of MCCase from *P. aeruginosa*

MCCase homologous carboxylases have been described as multimers containing α and β subunits that are essential for their function (Jitrapakdee and Wallace 2003). Protein co-expression is a convenient technique commonly used in the study of multimeric enzymes, because the subunits of some hetero-oligomeric proteins need to be expressed at the same time to favor the correct folding to produce a functional enzyme complex and avoid the formation of insoluble protein aggregates or non-active complexes (Kerrigan et al. 2011), as has been shown for the E and S subunits of the ornithine aminomutase (Chen et al. 2004). For this purpose, the *liuD* and *liuB* genes that encode for α and β subunits of the MCCase from *P. aeruginosa* were cloned into the pCDF-Duet1 vector, producing the pCDB3 co-expression plasmid. This strategy added a 6-histidine-tag and an S-tag epitopes into α and β subunits of the MCCase, respectively (Fig. S1a, supplementary material). These tags were used for the subsequent identification of each subunit (Terpe 2003). In addition, the His-tag incorporated in the LiuD recombinant protein allowed us to perform the one-step purification of the active complex of the MCCase.

Co-expression, purification, and immunodetection of the MCCase from *P. aeruginosa*

Escherichia coli BL21 competent cells were transformed with the pCDB3 DNA plasmid to induce expression and achieve purification of the MCCase. Different times and temperatures of incubation were tested for the expression of the MCCase. A slightly predominant band was observed after 8 and 12 h of the induction with IPTG when analyzing the cell extracts in SDS-PAGE. The molecular weight (MW) determined for this band corresponded to that calculated for the β subunit of the MCCase (data not shown). Therefore, we proceeded to purify the MCCase by affinity chromatography using the His-tag protein fusion

with the α subunit (LiuD-His protein). With this procedure, three bands were displayed in the SDS-PAGE gel, one corresponded to the theoretical MW for the α subunit (74 kDa) and two bands were near to the MW of the β subunit (around 60 kDa; Fig. 2a). To verify the identity of these proteins an immunodetection analysis was performed. When the avidin-HRP conjugated probe (commonly used for biotinylated proteins identification) was used against the obtained purified MCCase, a band of 74 kDa was developed, confirming the presence of the α -subunit of the MCCase (biotinylated subunit; Fig. 2b). To identify the β -subunit, the anti-LiuB polyclonal antibody was used, which revealed two bands with a MW around 60 kDa, suggesting that these proteins could correspond to the β -subunit of the MCCase (Fig. 2c). In addition, the co-expressed β -subunit was identified using an probe against the S-tag; it was observed that only the upper band was recognized by the probe (Fig. 2d). Results of immunodetection indicated that the two protein bands identified around 60 kDa, correspond to β -subunit of the MCCase and suggesting that the smaller protein band obtained after the co-purification procedure corresponded to the β -subunit without the S-tag epitope, which may be produced due to proteolysis.

To determine the MCCase subunits ratio, a densitometric analysis of the SDS-PAGE was conducted. It was observed that the intensity of the band corresponding to the α -subunit was equal to the sum of the intensities of the two bands of the β subunit, suggesting a 1:1 (α : β) expression ratio for the MCCase.

Determination of the oligomerization state of the MCCase from *P. aeruginosa*

To estimate the oligomerization state of the active MCCase complex, the enzyme was subjected to gel filtration chromatography analysis. A MW of 490 kDa was determined for the MCCase complex (Fig. 3a). Furthermore, densitometric results of the MCCase subunits separated by SDS-PAGE indicated that the subunits ratio in the oligomer was 1:1, suggesting that the MCCase complex is constituted as a heterotetramer ($\alpha\beta$)₂, as shown for the MCCases from *P. citronellolis*, *Achromobacter* sp., *P. sativum*, and *S. tuberosum* (Fall and Hector 1977). Interestingly, the purified MCCase was active only for about 24 h, so we proceeded to verify whether a change in the oligomerization state of MCCase was related to the loss of enzymatic activity using filtration chromatography. When the freshly purified active MCCase was assayed, it appeared as a single peak on the gel filtration profile (Fig. 3b), however when the MCCase was stored for 3 days, at 4°C, it appeared as two peaks in the chromatographic profile, corresponding to lower and higher MW than that of the

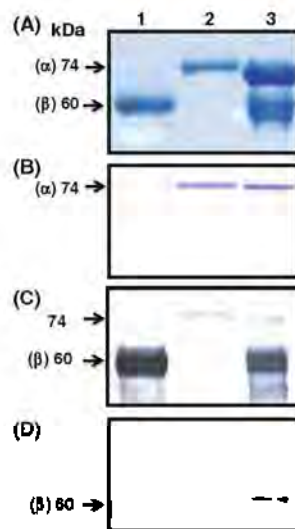


Fig. 2 Purification and western-blot analysis of the co-expressed and co-purified MCCase from *P. aeruginosa*. Purified recombinant α and β subunits of the MCCase were the same throughout a–d. Lanes: 1, MCCase β -subunit (LiuB-His) expressed and purified independently; 2, MCCase α -subunit (LiuD-His) expressed and purified independently; 3, co-expression and co-purification of the MCCase enzyme (α subunit (LiuD-His) and β subunits (LiuB-S)) using the LiuD-His subunit as protein-resin binding. a Coomassie blue-stained SDS-PAGE analysis. b Western-blot detection using avidin-HRP conjugate as probe to identify biotinylated proteins. c Western blot analysis of MCCase subunits using polyclonal anti LiuB-His as the first antibody and anti-mouse HRP as the secondary antibody. d Western blot analysis of MCCase using the anti-S-Tag-HRP conjugate as probe. Protein molecular weight markers in kDa are indicated as arrows

active enzyme (Fig. 3b). This result was more evident after 5 days of enzyme storage, when it was observed that the high MW species predominated over the lower MW species (Fig. 3b). This chromatographic pattern could be due to the disaggregation of the native $(\alpha\beta)_4$ MCCase, probably producing oligomers of $(\alpha\beta)_2$, $(\alpha\beta)_3$ or $(\alpha\beta)_6$, and the subsequent aggregation of the MCCase in inactive forms.

Enzymatic characterization of the co-expressed MCCase from *P. aeruginosa*

The enzymatic activity of the MCCase was measured through the carboxylation of the MC-CoA by radiolabeling assay (Hector and Fall 1976). The optimal temperature and pH shown for the MCCase were 30°C and 9.0, respectively (Fig. S1b, supplementary material). At optimal conditions, the MCCase showed a hyperbolic kinetic behavior, corresponding to a Michaelian enzyme (Fig. 4). In the Lineweaver–Burk plot, a straight line was obtained

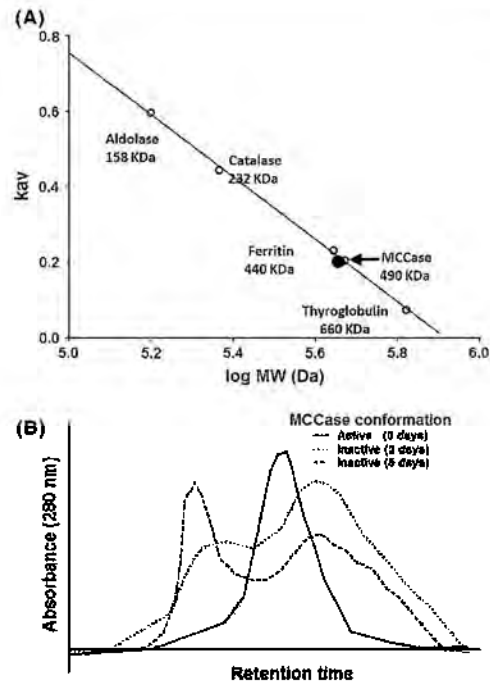


Fig. 3 Determination of the molecular weight and oligomerization state of the MCCase from *P. aeruginosa*. a Gel filtration chromatography was used to determine the native molecular weight of the MCCase complex (black point). The r^2 value of the linear regression of the standards was 0.99. b Gel filtration chromatography profile of the MCCase freshly purified and after 3 and 5 days of storage

confirming the Michaelian behavior (Fig. 4, see inset). The non-linear regression analysis rendered values of K_m for its substrate and V_{max} of 168 μM and 430 $\text{nmoles mg}^{-1}\text{min}^{-1}$ for the MC-CoA, respectively; similar results were described for MCCases from *P. citronellolis* (Fall 1981), plants (Baldet et al. 1992; Alban et al. 1993), and animals (Lau et al. 1980; Chu and Cheng 2007).

Previously, the kinetic behavior of the *P. aeruginosa* MCCase was described through an enzyme preparation in which the α and β subunits were expressed and purified separately and then the complex was reconstituted. In that study, the MCCase showed sigmoidal kinetics, suggesting a cooperative enzyme behavior with a $K_{0.5}$ of 9.8 μM for its substrate and a V_{max} of 279 $\text{nmoles mg}^{-1}\text{min}^{-1}$ (Aguilar et al. 2008). The difference in kinetic behavior of the enzymes obtained in both procedures could be explained as a result of the formation of a partially coupled MCCase complex in the case of the reconstituted enzyme, diminishing the V_{max} and interfering with the interaction between

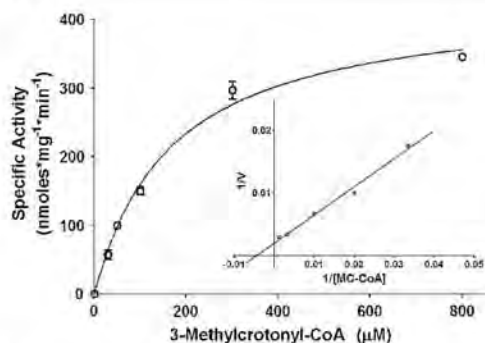


Fig. 4 Kinetic behavior of the co-expressed and co-purified MCCase from *P. aeruginosa*. Line shows the non-linear fitting to the Michaelis-Menten equation ($R^2 = 0.97$). Inset, Lineweaver-Burk plot of the data ($R^2 = 0.98$). Mean \pm SE of the data are indicated as error bars

subunits in the complex, thereby altering the binding of the substrate.

The GCCase, a nearly homologous enzyme of MCCase, was capable of using MC-CoA as alternative substrate in addition to G-CoA (its true substrate) (Aguilar et al. 2008); for this reason, the G-CoA was assayed as alternative substrate for the MCCase. The results with G-CoA indicated that the MCCase was unable to use this substrate as reported previously (Aguilar et al. 2008).

In summary, the results obtained indicate that the oligomerization state of the MCCase from *P. aeruginosa* is $(\alpha\beta)_4$ as described for homolog enzymes from other sources. In addition, this MCCase is specific for MC-CoA as substrate and unable to use G-CoA. The co-expression of both α and β subunits of the MCCase, as well as its co-purification using LiuD-His-tagged protein to obtain the MCCase complex in one step, is suitable to obtain an active biotin-carboxylase enzyme.

Acknowledgments We thank the *Pseudomonas* Genome Project and the *Pseudomonas aeruginosa* Community Annotation Project. This research was funded by CONACYT (83192, 106567) and C.I.C. 2.14/UMSNH grants.

References

- Aguilar JA, Zavala AN, Díaz-Pérez C, Cervantes C, Díaz-Pérez AL, Campos-García J (2006) The *atu* and *liu* clusters are involved in the catabolic pathways for acyclic monoterpenes and leucine in *Pseudomonas aeruginosa*. *Appl Environ Microbiol* 72(3):2070–2079. doi:10.1128/AEM.72.3.2070-2079.2006
- Aguilar JA, Díaz-Pérez C, Díaz-Pérez AL, Rodríguez-Zavala JS, Nikolau BJ, Campos-García J (2008) Substrate specificity of the 3-methylcrotonyl coenzyme A (CoA) and geranyl-CoA carboxylases from *Pseudomonas aeruginosa*. *J Bacteriol* 190(14):4888–4893. doi:10.1128/JB.00454-08
- Alban C, Baldet P, Axiotis S, Douce R (1993) Purification and Characterization of 3-Methylcrotonyl-Coenzyme A Carboxylase from Higher Plant Mitochondria. *Plant Physiol* 102(3):957–965. doi:10.1104/pp.102.3.957
- Baldet P, Alban C, Axiotis S, Douce R (1992) Characterization of biotin and 3-methylcrotonyl-Coenzyme A carboxylase in higher plant mitochondria. *Plant Physiol* 99(2):450–455. doi:10.1104/pp.99.2.450
- Campos-García J (2010) Metabolism of acyclic terpenes by *Pseudomonas*. In: Ramos JL, Filloux A (eds) *Pseudomonas*. Springer, Netherlands, pp 235–253. doi:10.1007/978-90-481-3909-5_8
- Chávez-Avilés M, Díaz-Pérez AL, Reyes-de la Cruz H, Campos-García J (2009) The *Pseudomonas aeruginosa* *liuE* gene encodes the 3-hydroxy-3-methylglutaryl coenzyme A lyase, involved in leucine and acyclic terpene catabolism. *FEMS Microbiol Lett* 296:117–123. doi:10.1111/j.1574-6968.2009.01624.x
- Chen HP, Hsui FC, Lin LY, Ren CT, Wu SH (2004) Coexpression, purification and characterization of the E and S subunits of coenzyme B(12) and B(6) dependent *Clostridium sticklandii* D-ornithine aminomutase in *Escherichia coli*. *Eur J Biochem* 271(21):4293–4297. doi:10.1111/j.1432-1033.2004.04369.x
- Chu CH, Cheng D (2007) Expression, purification, characterization of human 3-methylcrotonyl-CoA carboxylase (MCCC). *Protein Expr Purif* 53(2):421–427. doi:10.1016/j.pep.2007.01.012
- Díaz-Pérez AL, Zavala-Hernández AN, Cervantes C, Campos-García J (2004) The *gnyRDBHAL* cluster is involved in acyclic isoprenoid degradation in *Pseudomonas aeruginosa*. *Appl Environ Microbiol* 70(9):5102–5110. doi:10.1128/AEM.70.9.5102-5110.2004
- Diez TA, Wurtele ES, Nikolau BJ (1994) Purification and characterization of 3-methylcrotonyl-coenzyme-A carboxylase from leaves of *Zea mays*. *Arch Biochem Biophys* 310(1):64–75. doi:10.1006/abbi.1994.1141
- Fall RR (1981) 3-Methylcrotonyl-CoA and geranyl-CoA carboxylases from *Pseudomonas citromellolis*. *Methods Enzymol* 71(Pt C):791–799
- Fall RR, Hector ML (1977) Acyl-coenzyme A carboxylases. Homologous 3-methylcrotonyl-CoA and geranyl-CoA carboxylases from *Pseudomonas citromellolis*. *Biochemistry* 16(18):4000–4005
- Guan X, Diez T, Prasad TK, Nikolau BJ, Wurtele ES (1999) Geranyl-CoA carboxylase: a novel biotin-containing enzyme in plants. *Arch Biochem Biophys* 362(1):12–21. doi:10.1021/bi00637a010
- Hector ML, Fall RR (1976) Evidence for distinct 3-methylcrotonyl-CoA and geranyl-CoA carboxylases in *Pseudomonas citromellolis*. *Biochem Biophys Res Commun* 71(3):746–753. doi:10.1016/0006-291X(76)90894-9
- Hoschle B, Gnau V, Jendrosseck D (2005) Methylcrotonyl-CoA and geranyl-CoA carboxylases are involved in leucine/isovalerate utilization (*Liu*) and acyclic terpene utilization (*Atu*), and are encoded by *liuB/liuD* and *atuC/atuF*, in *Pseudomonas aeruginosa*. *Microbiology* 151(Pt 11):3649–3656. doi:10.1099/mic.0.28260.0
- Jitrapakdee S, Wallace JC (2003) The biotin enzyme family: conserved structural motifs and domain rearrangements. *Curr Protein Pept Sci* 4(3):217–229. doi:10.2174/1389203033487199
- Kerrigan JJ, Xie Q, Ames RS, Lu Q (2011) Production of protein complexes via co-expression. *Protein Expr Purif* 75(1):1–14. doi:10.1016/j.pep.2010.07.015



- Lau EP, Cochran BC, Fall RR (1980) Isolation of 3-methylcrotonyl-coenzyme A carboxylase from bovine kidney. *Arch Biochem Biophys* 205(2):352–359. doi:10.1016/0003-9861(80)90117-4
- Lipmann F, Tuttle CL (1945) A specific micromethod for the determination of acyl phosphates. *J Biol Chem* 159:21–28
- Lowry OH, Rosebrough NJ, Farr AL, Randall RJ (1951) Protein measurement with the Folin phenol reagent. *J Biol Chem* 193(1):265–275
- Sambrook J, MacCallum P, Russell D (2001) *Molecular cloning: a laboratory manual*, 3rd edn. CSH press, New York
- Schiele U, Niedermeier R, Sturzer M, Lynen F (1975) Investigations of the structure of 3-methylcrotonyl-CoA carboxylase from *Achromobacter*. *Eur J Biochem* 60(1):259–266. doi:10.1111/j.1432-1033.1975.tb20998.x
- Terpe K (2003) Overview of tag protein fusions: from molecular and biochemical fundamentals to commercial systems. *Appl Microbiol Biotechnol* 60(5):523–533. doi:10.1007/s00253-002-1158-6
- Würtele ES, Nikolau BJ (2000) Characterization of 3-methylcrotonyl-CoA carboxylase from plants. *Methods Enzymol* 324:280–292. doi:10.1016/S0076-6879(00)24238-9

Supplementary material

[Click here to download Supplementary material: FigS1.ppt](#)

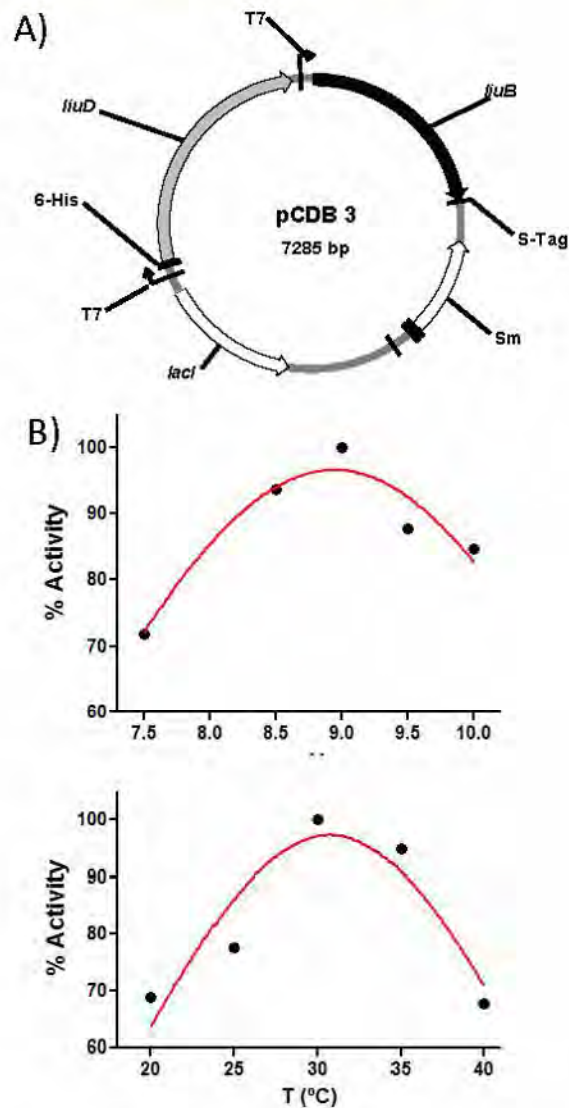


Figure S1.- A) Co-expression plasmid pCDB3. The plasmid contains both subunits of the MCCase from *P. aeruginosa* under the promoter T7 control. The α -subunit and β -subunit encoded by *liuD* and *liuB* genes were fused with 6-His and S-Tag peptides, respectively. B) Temperature and pH behavior for the MCCase recombinant enzyme.



Capitulo III:

“Structure modeling, docking and site-directed mutagenesis to identify residues of the active site of the β -subunit of 3-Methylcrotonyl-Coenzyme A carboxylase from *P. aeruginosa*: residues Ser187 and Arg51 are involved for substrate recognition in the MCCase”

VII.3) Capítulo III:

“Structure modeling, docking and site-directed mutagenesis to identify residues of the active site of the β -subunit of 3-Methylcrotonyl-Coenzyme A carboxylase from *Pseudomonas aeruginosa*: residues Ser187 and Arg51 are involved for substrate recognition in the MCCase”

Díaz-Pérez César¹, Rodríguez-Zavala José Salud², Díaz-Pérez Alma-Laura¹ and Campos-García Jesús¹. Enviado para su publicación a BMC, Microbiology.

VII.3.a) INTRODUCCIÓN

La 3-metilcrotonil-CoA carboxilasa (MCCasa) de *P. aeruginosa* es una enzima esencial en las rutas de degradación de leucina y monoterpenos acíclicos, la cual cataliza la carboxilación del 3-metilcrotonil-CoA para producir 3-metilglutaconil-CoA. La MCCasa pertenece a la familia de carboxilasas dependientes de biotina (BDC). En esta familia, el sitio de unión del acil-CoA se forma en la interface de un dímero de subunidades β . En este estudio se realizó el modelaje estructural, anclaje molecular y alineamientos de secuencias para identificar los residuos relacionados con el reconocimiento del substrato. Además, se efectuaron mutaciones sitio-dirigidas para confirmar las predicciones. El modelo molecular mostró que la MCCasa está formada por dos dominios de crotonasa, lo cual sugiere que la enzima evolucionó de un evento de duplicación génica. Alineamientos estructurales del modelo con las proteínas cristalizadas de la familia BDC indicó que el sitio activo está formado por dos zonas, y se encuentra localizado en la interface del dímero de subunidades β . El análisis de anclaje molecular con el substrato y el cofactor biotina sugiere que los residuos R51 y S187, localizados en la zona I, están involucrados en la unión y orientación del substrato, mientras los residuos F417, Y422 y G423, localizados en la zona II, están involucrados en la carboxilación. La mutagénesis sitio-dirigida mostró que los cambios Y422D y S187A abaten por completo la actividad enzimática, mientras que los cambios R51A, F417Y y G423A modifican los parámetros cinéticos de la enzima. Estos resultados sugieren que las dos zonas son esenciales para la actividad de la MCCasa, la zona I es importante para el reconocimiento del substrato, mientras que la zona II ayuda en la reacción de carboxilación.

ABSTRACT

The enzyme 3-methylcrotonyl-CoA carboxylase (MCCase) from *Pseudomonas aeruginosa* is essential for the assimilation of leucine and acyclic monoterpenes, catalyzing the carboxylation of 3-methylcrotonyl-CoA (MC-CoA) to 3-methylglutaconyl-CoA. The MCCase belongs to the biotin-dependent carboxylases, in which the binding site of the substrate acyl-CoA is formed in the interface of the β -subunit dimers. In this study homology modeling, molecular docking and sequence alignment analyses were performed to predict the residues involved in substrate recognition, and site directed mutagenesis was carried out to confirm the predictions. The structural model of MCCase showed that it is conformed for two crotonase domains, suggesting that the enzyme evolved from a common origin by a domain duplication event. Structure alignment of the model with crystal structures of biotin-dependent carboxylases (BDC) family, indicated that the active site is constituted by two zones, and is located in the interface of the β -subunit dimer. Docking analysis of the MCCase with the substrate and biotin cofactor suggests that residues R51 and S187 are involved in the binding and orientation of the substrate (located in zone I), whereas the residues F417, Y422 and G423 with carboxylation (located in zone II). Site-directed mutagenesis showed that the change Y422D and S187A completely abolished the enzyme activity, whereas the changes in residues R51A, F417Y and G423A modified the kinetics parameters of the enzyme. The results suggest that both zones are essential for the MCCase activity: the zone I in the MCCase is important for substrate recognition, while that the zone II for the reaction of carboxylation.

KEYWORDS

Pseudomonas aeruginosa, biotin-carboxylases, 3-methylcrotonyl-CoA, molecular modeling, site-directed mutagenesis.

INTRODUCTION

Pseudomonas aeruginosa is a bacterium capable to utilize branched-chain amino acids such as leucine as sole carbon source, through the leucine/isovalerate assimilation pathway that has been studied in detail both, in prokaryotes and eukaryotes. The route is composed by the upper activation and the central degradation pathways (Campos-García, 2010). The central part is known as leucine/isovalerate utilization pathway (LIU). It has a dual role in the *P. aeruginosa* catabolism, since the central LIU route is essential for the acyclic terpenes utilization pathway (ATU) (Fig. 1A; Aguilar *et al.*, 2006; Díaz-Pérez *et al.*, 2004). These pathways converge at the 3-methylcrotonyl-CoA carboxylase (MCCase, EC 6.4.1.4) that catalyzes the carboxylation of the substrate 3-methylcrotonyl-CoA (MC-CoA) to produce 3-methylglutaconyl-CoA (MG-CoA) (Aguilar *et al.*, 2008).

The *P. aeruginosa* MCCase is constituted by two subunits (α and β) forming an $(\alpha\beta)_4$ active complex (Diaz-Perez *et al.*, in press), both α and β subunits are essential for enzyme function (Aguilar *et al.*, 2006). However recently the crystal structure of the the MCCase has been resolved indicating that the enzyme is conformed as an $(\alpha\beta)_6$ complex (Huang *et al.*, 2011). The α -subunit is encoded by the *liuD* gene (PA2012), it carries the cofactor binding domain (BCCP) and the biotin carboxylation domain (BC) (Fig. 1B). The biotin cofactor is covalently bound to a lysine residue. The β -subunit, encoded by the *liuB* gene (PA2014), harbors the transcarboxylase domain (TC). The carboxylation reaction catalyzed by MCCase involves both subunits and two catalytic steps (Fig. 1B). In the first step, the biotin molecule is attached to BCCP domain and it is carboxylated to produce carboxy-biotin, this reaction is catalyzed by the BC domain from α -subunit. In the second reaction, catalyzed by TC domain from the β -subunit, the carboxyl group is transferred from the biotin to the substrate MC-CoA, producing the MG-CoA (Aguilar *et al.*, 2008; Campos-García, 2010; Huang *et al.*, 2011).

Phylogenetically, the MCCase belongs to the class I of the biotin-dependent carboxylases (BDC), which use biotin as cofactor. In this protein family, the BC and

BCCP domains have conserved amino acid sequences and structural similarities (Jitrapakdee and Wallace, 2003). On the other hand, members of the BDC family share sequence and structural similarities in the TC domain, but there are feature variations of this area depending on the substrate used by the enzyme, either an acyl-CoA or pyruvate (Toh *et al.*, 1993; Huang *et al.*, 2011). For example, the acetyl-CoA (ACCase) and propionyl-CoA carboxylases (PCCase) enzymes from *Streptomyces coelicolor* use acetyl-CoA and propionyl-CoA as substrates, respectively, sharing the same biotinylated subunit (BC and BCCP domains), but the TC domains is different (Diacovich *et al.*, 2004). The subunits of the BDCs' family that harbor the TC-domain perform their function as dimers, in a head-to-tail arrangement of two monomers. The active sites of the CT-dimer were located in cavities at the interface of the dimer, with contribution from each monomer to form the active-site (Diacovich *et al.*, 2004; Hall *et al.*, 2003; Huang *et al.*, 2010; Zhang *et al.*, 2003; Huang *et al.*, 2011). Previous studies with the acyl-CoA carboxylase from *S. coelicolor* showed that the residue D422 belongs to the active site and it is important for substrates discrimination (Diacovich *et al.*, 2004). However in the PCCase from *Ruegeria pomeroyi*, the mutation of the equivalent residue did not alter the substrate preference, suggesting that although the active site of these enzymes is conserved, the role of residues may be different among members of this family (Huang *et al.*, 2010; 2011).

The biochemical characterization and expression studies of MCCase from *P. aeruginosa* related to leucine/acyclic terpenes metabolism have been achieved (Aguilar *et al.*, 2006; Aguilar *et al.*, 2008), and recently, the MCCase from *P. aeruginosa* has been used for crystallographic studies, to understand the human disease associated with accumulated mutations in patients that shows methylcrotonylglycinuria (Huang *et al.*, 2011). Nevertheless, the structure and active-site residues of the enzyme remain to be characterized. In this study a homology model, molecular docking and sequence alignment analyses were performed to predict the residues involved in substrate recognition and site directed mutagenesis was carried out to confirm the predictions.

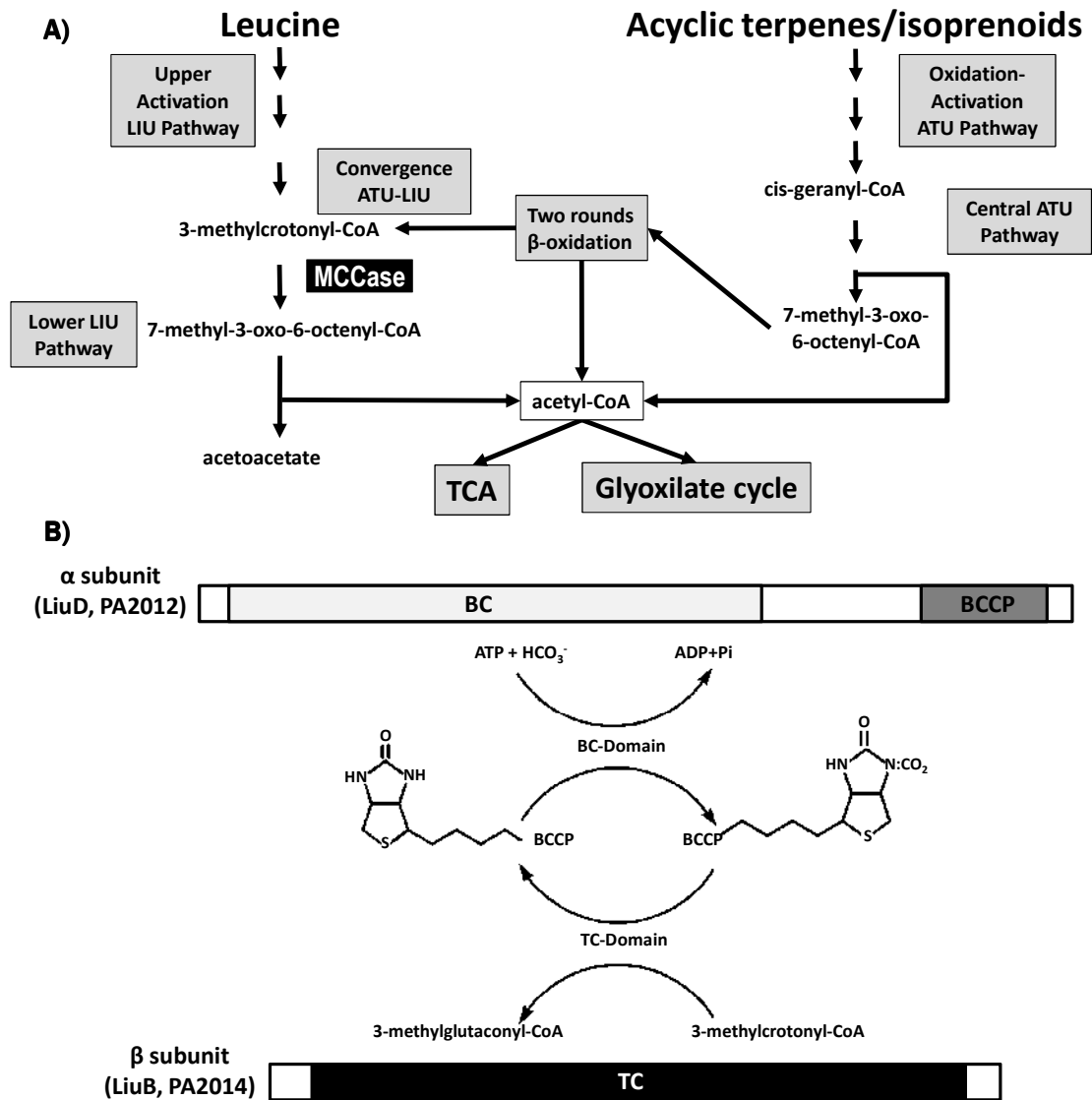


Figure 1.- Metabolic role and subunits of the 3-methyl-crotonyl-CoA carboxylase from *P. aeruginosa* PAO1. A) General steps that comprise LIU and ATU metabolic pathways involved in leucine and acyclic terpenes assimilation in *P. aeruginosa*. MCCase is the convergent enzyme for both catabolic pathways (Adapted from Campos-Garcia, 2010). MCCase is indicated in a black square. B) The chemical reaction catalyzed by MCCase involves two steps carried out for each subunit. In the first step, the biotin-cofactor, covalent bound to BCCP domain through an K residue, is carboxylated in the α subunit. The second step consists in the transference of carboxyl group to the MC-CoA substrate occurring in the TC domain from the β subunit.

MATERIALS AND METHODS

Bacterial strains, plasmids and culture conditions

The bacterial strains and plasmid used in this work are listed in Table S1 (supplementary material). The strains were grown in Luria-Bertani broth (LB) or 2XYT expression-medium (tryptone 16 g/L, yeast extract 10 g/L, NaCl 5 g/L) at 30°C and shaking. Solid media were prepared by adding 1.5% agar. Antibiotics used were streptomycin at 30 µg/ml and chloramphenicol at 30 µg/ml.

Site-directed mutagenesis

Site-directed mutagenesis was carried out by standard PCR method, using GeneTailor Site-Directed Mutagenesis System (Invitrogen). The PCR amplification was performed according to manufacturer's instructions, using the pCYC-B plasmid as pDNA template. The primers used to introduce the mutations into *LiuB* gene are listed in Table S2 (supplementary material). The PCR-obtained plasmids were transformed into the *E. coli* strain (MAX Efficiency DH5α-T1 chemically competent cells; Invitrogen). The plasmids containing the *liuB* mutant genes were purified using the Wizard Plus SV Minipreps DNA Purification System (Promega) and subjected to DNA sequencing to confirm the mutation. The cloned or mutated genes were excised by *Bsu36I* and *HindIII* endonucleases and subcloned into the pCBD3 dual co-expression plasmid (Diaz-Perez *et al.*, in press) rendering the pCBDMut plasmids, which were then transformed in *E. coli* JM101.

Expression and purification of recombinant MCCase

The pCBDMut plasmids containing the wild-type *liuD* gene and the respective mutated *liuB* genes were transferred into *E. coli* BL21 strain. The bacterial cells were cultured on 400 ml of 2XYT medium for 4 h at 37°C with shaking. Protein expression was induced by adding 0.1 mM of isopropyl-β-D-thiogalactopyranoside (IPTG) and then the cultures were incubated for 12 h at 37°C with shaking. Cells were harvested by centrifugation at 5000 × *g* for 10 min at 4°C. The bacterial pellet was suspended in 10 ml of buffer A (50 mM Tris-HCl, 500 mM NaCl, pH 7.5) and disrupted by sonication

at 4°C. Cell extract was centrifuged for 30 min at 15,000 × *g* at 4°C to eliminate undisrupted cells and cell debris, obtaining the cells-free crude extracts. Recombinant proteins were purified from cells-free crude extracts according to the TALON™ purification kit protocol (DB-Biosciences). The resin-protein mixture 250 µl was washed four times with 1 ml of wash buffer (buffer A plus 20 mM imidazole) and then with 500 µl of elution buffer (buffer A containing 250 mM imidazole) to elute the protein. Protein purification for the MCCases containing both α and β subunits were achieved with 95% purity (Fig. S1, supplementary material). The purified MCCase enzyme was diluted 1:4 (v/v) with buffer A plus 50% (v/v) glycerol. Protein concentration was determined by the Lowry method (Lowry *et al.* 1951). The native molecular weight of the purified MCCase was determined using an 80 cm-long column packed with Sephacryl-S-300 as described (Díaz-Pérez *et al.*, in press).

G-CoA synthesis and Carboxylase activity determination

Synthesis of geranyl-CoA (G-CoA) was reached by the mixed-anhydride method as described (Aguilar *et al.*, 2008). MC-CoA and ATP were supplied by SIGMA.

MCCase and GCCase activities of the purified proteins were measured by the incorporation of ¹⁴CO₂ into acidic nonvolatile material as previously described (Aguilar *et al.*, 2008, Díaz-Pérez *et al.*, in press).

Molecular modeling

A Blast search against the sequences of Protein Data Bank (PDB) was carried out, selecting three crystallized protein templates for more studies. Multiple sequence alignment of the β-subunit of the MCCase of *P. aeruginosa* with the templates was performed using T-COFFE software (Notredame *et al.*, 2000). The three-dimensional homology modeling of the β-subunit of the *P. aeruginosa* MCCase and the LiuB site-directed mutants was carried out using Modeller 9v7 (Fiser and Sali, 2003). A recursive protocol was implemented to obtain models with good quality. The quality of models was checked employing Procheck 3.5 (Laskowski *et al.*, 1993), VERIFY3D (Luthy *et al.*, 1992), and ProSA (Wiederstein and Sippl, 2007) programs. The root-mean-square deviation (RMSD) values between superimposed models and



crystallographic structure were calculated by using SuperPose software (Maiti *et al.*, 2004). Volume of the active site was calculated with 3V server (Voss and Gerstein, 2010). All figures were generated with PyMOL Molecular Graphics System, Version 1.2r1, Schrödinger, LLC.

Molecular docking analysis

The M-CoA and biotin-lysine substrates were docked into the active site of the refined model from the MCCase LiuB structure using AutoDock Vina (Trott and Olson, 2010) and AutoDock 4.0 (Morris *et al.*, 2009) programs. The MC-CoA and biotin-lysine molecules were modeled with the VI EWDD molecular editor (Rome Center for Molecular Design, Lab) and used as ligands. The grid box for MC-CoA was set to cover the whole active site. Docking was run (100 times) using the Lamarckian genetic algorithm of AutoDock 4.0. The composite file for all possible conformers was analyzed with AutoDock Tools 5 (Sanner, 1999). The substrate conformation selected was in agreement with the conformation of the substrates from the crystallized protein structures used as templates. The figures were generated with PyMOL Molecular Graphics System, Version 1.2r1, Schrödinger, LLC.

RESULTS AND DISCUSSION

Molecular Modeling

To model the structure of the β -subunit of MCCase, suitable templates were obtained by Blast-searching against PDB database using as query the LiuB amino acids sequence (GenProt: NP_250704). The three best hits were selected and showed sequence identity higher than 30% (50% similarity) and sequence coverage higher than 90%. These templates were the β -subunits of propionyl-CoA carboxylase from *Streptomyces coelicolor* (1XNY) and *Thermotoga maritime* (1VRG), and subunit 12S of the transcarboxylase from *Propionibacterium freudenreichii* (1ON3).

Previously, the binding site of the substrate of the crystallized members of BDC was found to be formed in the interface of two β -subunits (Diacovich *et al.*, 2004; Hall *et al.*, 2003; Huang *et al.*, 2010; Zhang *et al.*, 2003). Additionally, it has been reported that the minimal active complex assembly for MCCase from *P. aeruginosa* was $(\alpha\beta)_4$ (Diaz-Perez *et al.*, in press). Furthermore, in all the crystallized carboxylases of the family, the β -subunit is involved in substrate recognition. Therefore, a homology tridimensional model of the *P. aeruginosa* MCCase was built using the β -subunit dimer of crystallized carboxylases. To verify the accuracy of the modeled structure, the model was superimposed with the crystal structures reported for the three templates (1XNY, 1VRG, and 1ON3). The RMSD values obtained by the alignments were 0.578, 0.456 and 0.495 Å for the templates 1ON3, 1VRG and 1XNY, respectively. These results indicated that the MCCase molecular model had a similar fold compared with the templates used (Fig. S2, supplementary material).

After three modeling rounds a tridimensional model with satisfactory parameters was obtained as verified with three different programs. The PROCHECK software was applied to qualify the combinations of ϕ and ψ residues angles in available zones of Ramachandran plot, Verify-3D was used to assess the quality of the model, and ProSA program to evaluate the energy of the structure using a distance based pair potential. The final model proposed was validated in terms of geometry and energy profiles using these programs (Table 1; Figs. S3-S5, supplementary material). Values

obtained indicate that the model had enough confidence to explore the protein–substrate interactions. The structural model obtained showed that each β -subunit monomer is made up by two well defined sub-domains which are intimately associated with each other, named here N- and C- domains (Fig. 2; a and a' domains or bn and b' domains). This two domains backbone fold is consistent with that of the crotonase/ClpP superfamily. Several proteins that contain a crotonase/ClpP domain are acyl-CoA–dependent enzymes (Hamed *et al.*, 2008), and this fold is also conserved in all BDC enzymes that use acyl-CoA as substrates (Holm and Rosenstrom, 2010). Also it was observed that the N- and C- domains which conformed the MCCase β -subunit share 25% sequence identity (42% positives). When the structures of N- and C- domains were superimposed a RMSD of 1.853 Å was obtained, indicating that their structures were similar. The main differences were found in the first 20 residues in both domains (Fig 2, see circles) and one insertion in each domain (located at residues 77-88 and 456-484). In addition, the RMSD value diminished to 1.253 Å when the insertions in the structures of N- and C- domains were excluded of the analysis (30% identity, 47% positives). Sequence similarity and structure arrangement suggest that the N- and C- domains of the MCCase β -subunit from *P. aeruginosa* have a common origin, and possibly, gene duplication involving a fusion event gave rise to the MCCase β -subunit.

In other context, the β -subunits of the BDC family assembly forming complexes of diverse oligomerization states. Examples of these oligomers are: dimers in the ACCase from yeast (Zhang *et al.*, 2003), hexamers in the bacterial PCCases (Diacovich *et al.*, 2004; Huang *et al.*, 2010) and bacterial transcarboxylase (Hall *et al.*, 2003). Differences in the oligomeric assembly among these enzymes is proposed to relay on the presence of an insertion of two helical motifs found in the C-domain of the β -subunit in the dimeric BDC. That insertion prevents the formation of complexes of higher molecular mass in the dimeric BDCs' conformers (Diacovich *et al.*, 2004; Huang *et al.*, 2010). The *P. aeruginosa* MCCase is conformed in its native/active conformation as a $\alpha_4\beta_4$ complex (Diaz-Perez *et al.*, in press), suggesting that the β -subunit forms a tetrameric structure. This conformation is feasible to occur because

the TC-domain of the MCCase β -subunit does not possess the amino acids insertion at the C-domain mentioned above, that avoids the formation of the tetrameric complex $\alpha_4\beta_4$ (Fig. 3). An additional insertion (20 amino acids) was found in the middle C-domain of the $\alpha_2\beta_2$ ACCase, forming a helix extension avoiding further oligomerization of the subunits; while in the $\alpha_4\beta_4$ MCCase model, the insertion is formed by 8 amino acids and in the $\alpha_6\beta_6$ PCCases this helix was not found (Fig. 3). These findings suggest that the length of the insertion is relevant to avoid bigger oligomerization states in the BDC family. However, recent crystallographic analyses of the MCCase of *P. aeruginosa* MCCase in the presence of coenzyme A showed an unanticipated architecture of $\alpha_6\beta_6$ (Huang *et al.*, 2011).

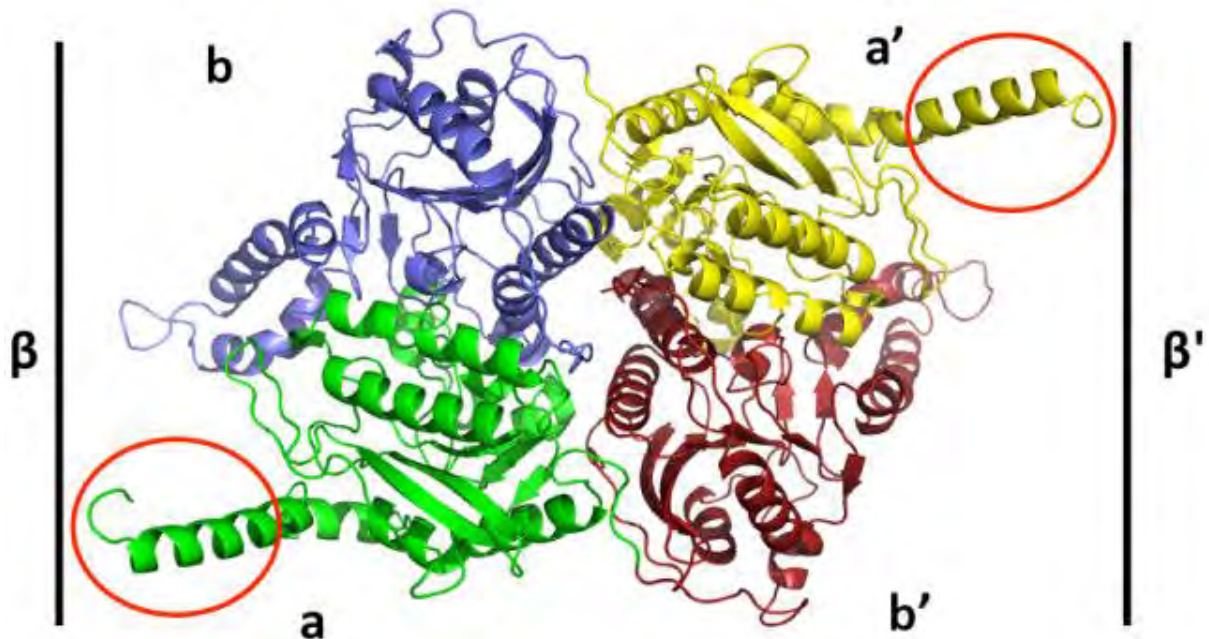


Figure 2.- Homology model of the MCCase β -subunit from *P. aeruginosa*. The functional unit of the MCCase is a dimer constituted for two crotonase motifs showed in different colors. Each monomer of the dimer, named here β and β' , is formed by two subdomains: a (green) and b (purple) domains for β monomer, and a' (yellow) and b' (cherry) domains for β' monomer. The first 20 residues in N-region of each monomer are marked with a red circle.

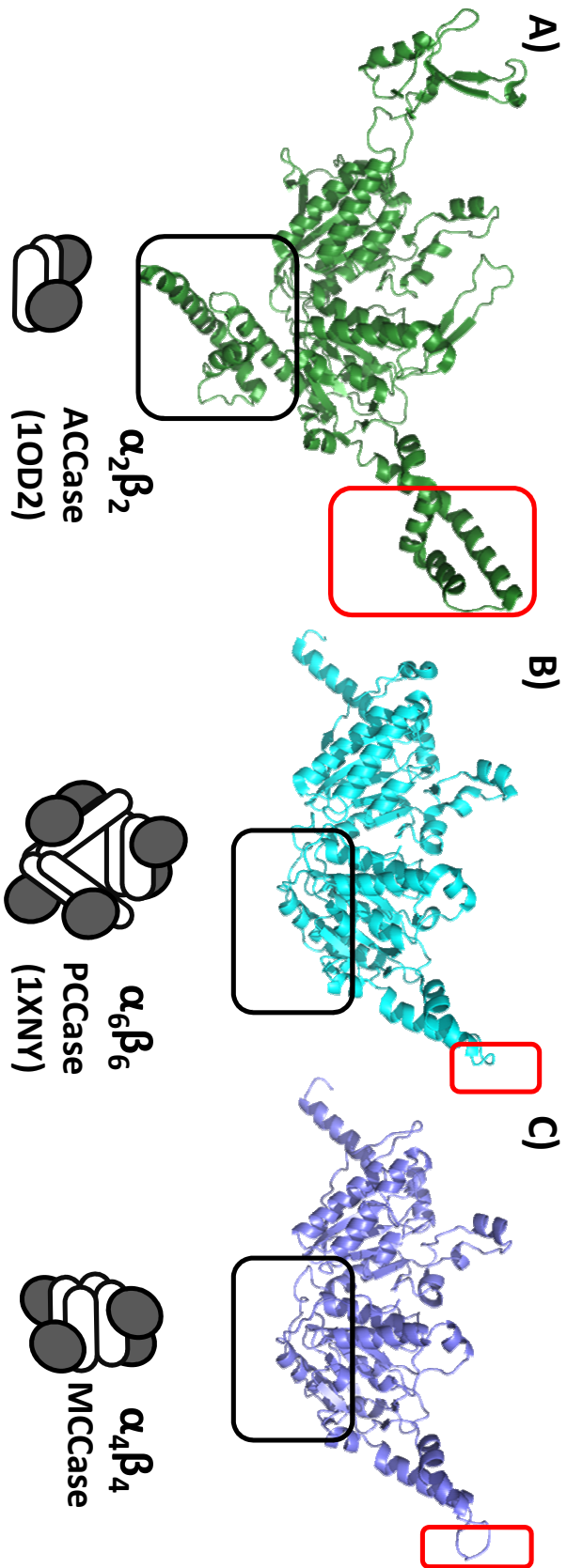


Figure 3. - Insertions of amino acids in subdomains of BDC prevent the formation of large complexes. A) ACCase from yeast has two insertions, and it only forms dimers. B) The PCCase from bacteria does not present these insertions and therefore it can form hexamers. C) The MCCase from *P. aeruginosa* under native conditions is active as a tetramer, and its structure only presents a small insertion. Amino acid insertions are indicated as rectangles.

Docking analysis

In order to study the interactions of the β -subunit MCCase from *P. aeruginosa* with its substrate, a molecular docking analysis with the validated model on the protein and the modeled substrate MC-CoA molecule was conducted. The TC-active site localization is conserved throughout the BDC protein family. The acyl-CoA binding pocket in the PCCase (Diacovich *et al.*, 2004; Huang *et al.*, 2010) and ACCase (Zhang *et al.*, 2003) is located at the interface between the N-domain of one monomer and the C-domain of its dimeric partner. Therefore, the active site of MCCase was located performing structural alignments with these resolved structures, and then, the docking analysis with the substrate was carried out. As expected, the alignment showed that the active site of the MCCase was located in the middle of a cavity formed by the dimer of β -subunits. The aliphatic chain of the substrate MC-CoA was positioned in the inner part of the cavity and the selected conformation of the substrate MC-CoA was in agreement with the conformation of the acyl-CoA from the templates used, rendering a binding energy of -8.2 kcal/mol (Fig. 4), similar to that obtained for propionyl-CoA when it was docked to PCCase structure (binding energy of -7.6 Kcal/mol).

After docking analysis the area of protein-substrate interaction was divided in two regions: zone I, where the residues interact with the Coenzyme A motif from the MCCase substrate, and zone II, where the interaction with the acyl-chain from MC-CoA occurs and the carboxylation reaction takes place (Fig. 4). In zone I, there is a patch of residues positively charged (R47, R51 and K53) that could interact with the phosphate groups of coenzyme A motif. According to docking results, the residue R51 may form two hydrogen bonds with the phosphate groups of coenzyme A (Fig. 4). The S187 is another residue located in this zone, and it could also form a hydrogen bond with the MC-CoA phosphate. These findings suggest that the substrate MC-CoA is attached to residue R51 and conducted to Zone II by the residue S187.

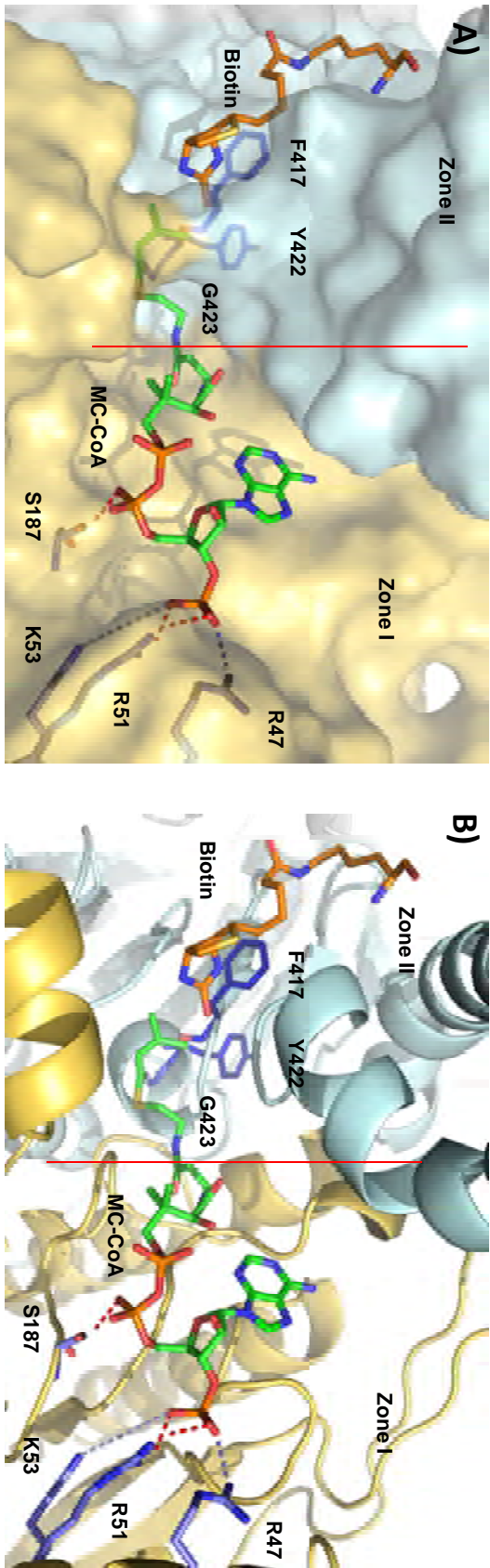


Figure 4.- Docking analysis of the substrate MC-CoA in the active site of the MCCase from *P. aeruginosa*. A) Surface conformation, B) Secondary structure conformation of the active site of the MCCase. The two zones in the active site are delimited by a dash line. The MCCase β and β' monomers are colored in yellow and blue respectively. Prediction shows the possible interactions of the substrate MC-CoA (green) with residues R51 and S187 in zone I of active site. Residues R47, R51, K53 and S187 are colored in blue. Biotin molecule is indicated in orange. Residues F417, Y422, and G423 are located in zone II of the active site, indicated in blue.

In the zone II, seventeen residues were located forming the substrate and biotin binding pockets (Fig. S6, supplementary material). To identify the residues that may be important in substrate stabilization in the zone II, a sequence alignment analysis with family members of this enzyme that have different substrate specificities was performed (Fig 5C). In zone II of the active site of MCCase, two variable residues were selected (Fig. 5), the F417 and G423 with respect to the GCCase paralogous protein (which has a preference for the substrate GC-CoA, but also can utilize MC-CoA as substrate) (Aguilar *et al.*, 2008). Both residues F417 and G423 could be responsible for substrate discrimination in the MCCase enzyme, in addition to residue Y422 that in the PCCase was found to be involved in substrate specificity (Diacovich *et al.*, 2004). Thus, to confirm that the substrate recognition is related to the less conserved residues, the amino acids F417, G423 and Y422 were selected for site-directed mutagenesis analysis.

Site-directed mutagenesis

Site directed mutagenesis was conducted with residues R51A and S187A which were changed to alanine, these residues located in zone I in the MCCase model, showed interactions with the phosphate groups of the substrate MC-CoA (Fig. 4). For mutagenesis in zone II, the residues replaced were F417Y, Y422D and G423A, respectively. F417Y and G423A changes were designed to validate whether these residues are responsible of the acyl-group discrimination between the substrates MC-CoA and GC-CoA as suggested in the model (Fig. 4). Y422D change was designed to determine whether this residue has relevance in substrate discrimination and catalytic properties, as it occurs in PCCase from *S. coelicolor* (Diacovich *et al.*, 2004). Additionally, a double mutant F417Y and G423A (namely FG mutant) was designed to resemble the amino acids sequence present in the active site of the enzyme GCCase enzyme, which although has the ability to use both GC-CoA and MC-CoA substrates, it has a preference for GC-CoA over MC-CoA (Fig. 5C).

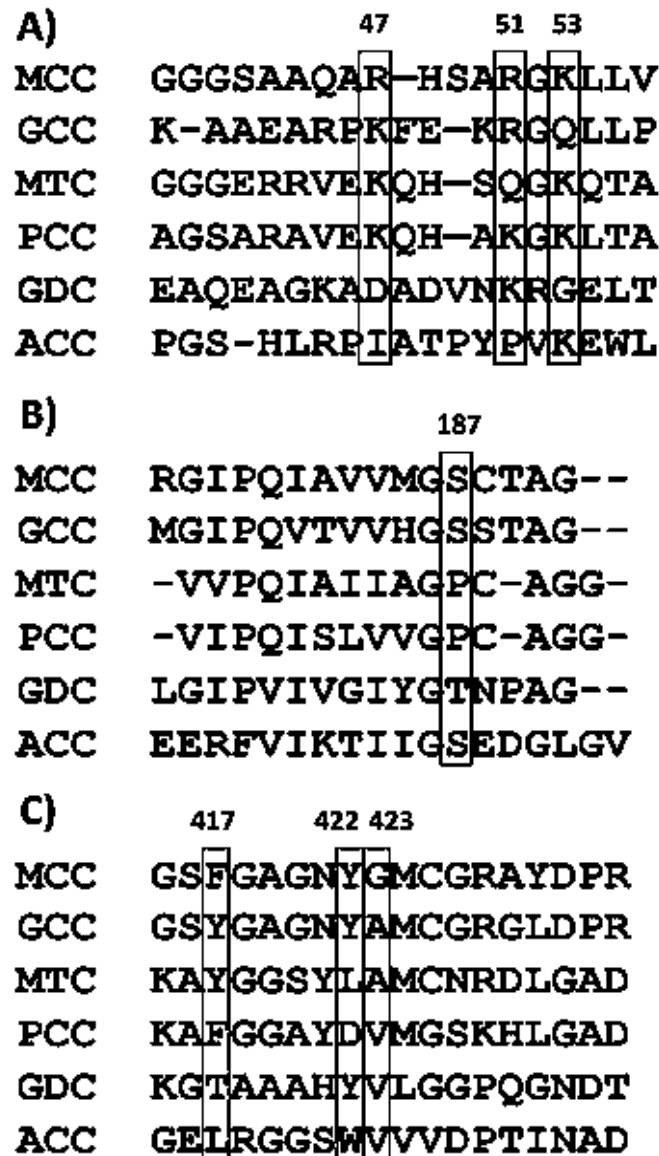


Figure 5.- Sequence alignment of the zones I and II of the MC-CoA-binding site of MCCase. A) Sequence alignment of the residues around R51 in zone I. B) Sequence alignment of the residues around S187 in zone I. C) Sequence alignment of the residues mutated in Zone II. Descriptions and GeneProt ID are as follow. MCC: β -subunit of Methylcrotonyl-CoA carboxylase from *P. aeruginosa*, NP_250704; GCC: β -subunit of Geranyl-CoA carboxylase from *P. aeruginosa*, NP_251578; MTC: 12S subunit of Methylmalonyl-CoA carboxyltransferase from *P. freudenreichii*, Q8GBW6; PCC: β -subunit of propionyl-CoA carboxylase from *S. coelicolor*, NP_629079; GDC: α -subunit of Glutaconyl-CoA decarboxylase from *Acidaminococcus fermentans*, Q06700; ACC: Acetyl-CoA carboxylase from *Saccharomyces cerevisiae*, AAA20073. The numbers and boxes over the alignments remark the residues mutated in this work.

Mutants of MCCase β -subunit were sub-cloned into a co-expression vector with the cloned MCCase α -subunit for co-expression and purification as previously designed (Diaz-Perez *et al.*, in press). The purified MCCase enzymes showed 95% purity as judged from SDS-PAGE (Fig. S1, supplementary material). The mutant S187A was inactive for carboxylation of the MC-CoA (Table 2), indicating that the S187 is essential for the catalysis. Docking analysis suggested that residue S187 and R51 stabilize the MC-CoA in the active site by hydrogen bonds formation. None other residues were found near to residue S187 that could play the role interacting with the substrate conferred for this residue. Docking analysis indicated that the mutant S187A was unable to accommodate the substrate in a suitable conformation for carboxylation (Fig. 6A).

The mutant R51A was active but it showed half of the activity compared with the wild type enzyme and also presented a value of K_m two-fold higher (Table 2). Residual activity in R51A mutant can be explained by the presence of residues R47 and K53 nearby residue R51, which could partially stabilize the substrate in the active site (Fig. 4). These results indicated that interaction of residue S187 is essential for the *P. aeruginosa* MCCase enzymatic activity, probably stabilizing the interaction with the substrate in a suitable position for catalysis, while the region of residue R51 and neighbor residues R47 and K53, conforms a structural pocket with positive changes suitable for the interaction with the coenzyme A motif (Fig. 4).

Mutagenesis experiments for residues located in the MCCase zone II (F417Y, Y422D and G423A changes) showed an increase on the activity of mutant F417Y and the double mutant FG of 2.9 and 1.5 folds, respectively, compared with the wild type enzyme (Table 2). The mutant G423A presented 30% less activity than wild type, while Y422D was completely inactive. Early studies with the *S. coelicolor* acyl-CoA carboxylase showed that the residue D422 (equivalent to Y422 in the MCCase) is important for discriminating between the substrates acetyl-CoA and propionyl-CoA (Diacovich *et al.*, 2004). However, in the MCCase the Y422D change caused the loss of function. In the model of MCCase, the residue Y422 was located at a distance of

6Å from the substrate, similar to PCCase from *R. pomeroyi* and chimeric PCCase from *Roseobacter denitrificans*, where equivalent mutations did not change the activity or substrate preference of the enzyme (Huang *et al.*, 2010). In the *P. aeruginosa* MCCase introducing an aspartic residue in position Y422 probably promotes a repulsion effect with the residues Q374 and S416, causing a structural change in the active site, and as a consequence, the substrate MC-CoA could not fit correctly into the active site (Fig. 6C). Additionally, an interaction between Y422 and D500 was not longer formed inducing the disruption of the active site. The analysis of the active site for the mutation Y422D in the model of MCCase confirmed this hypothesis, since with the mutation an increase on the volume of the cavity of zone II was observed, in comparison with the WT MCCase (Fig. 6D-E).

Replacement of G423A resulted in a reduction of 44% of the *K_m* and 32% in the *V_{max}* values, respect to the WT MCCase (Table 2). A possible explanation is that the structural perturbation introduced by the G423A mutation causes a minor change in local conformation that favored the MC-CoA-binding. However, the increase in affinity could also promote a delay in the release of the product from the active site, and as a consequence, a diminution of the MCCase activity was observed (Table 2).

The mutant F417Y showed a 2.9-fold increase in *V_{max}* and 2.3-fold increase in *K_m* constants, comparing with the WT MCCase (Table 2). Residue F417 in the model was orientated towards the biotin cavity; docking analysis of biotin molecule in the active site of the mutant F417Y showed that the hydroxyl group of residue tyrosine is feasible to interact with the molecule of biotin (Fig. 6F). The biotin role in the carboxylases is to transfer the carboxyl group to the substrate. The additional interaction produced for the mutation F417Y possibly promotes a diminution on the mobility of biotin in the active site of the MCCase, and although catalysis can take place whether the biotin release is delayed and the active site remains occupied for a longer time, the affinity for the substrate could be altered, as it was observed in the kinetic parameters for this mutant. These results suggest that in the *P. aeruginosa* MCCase the residues F417, Y422 and G423 constitute the active site. Additionally,

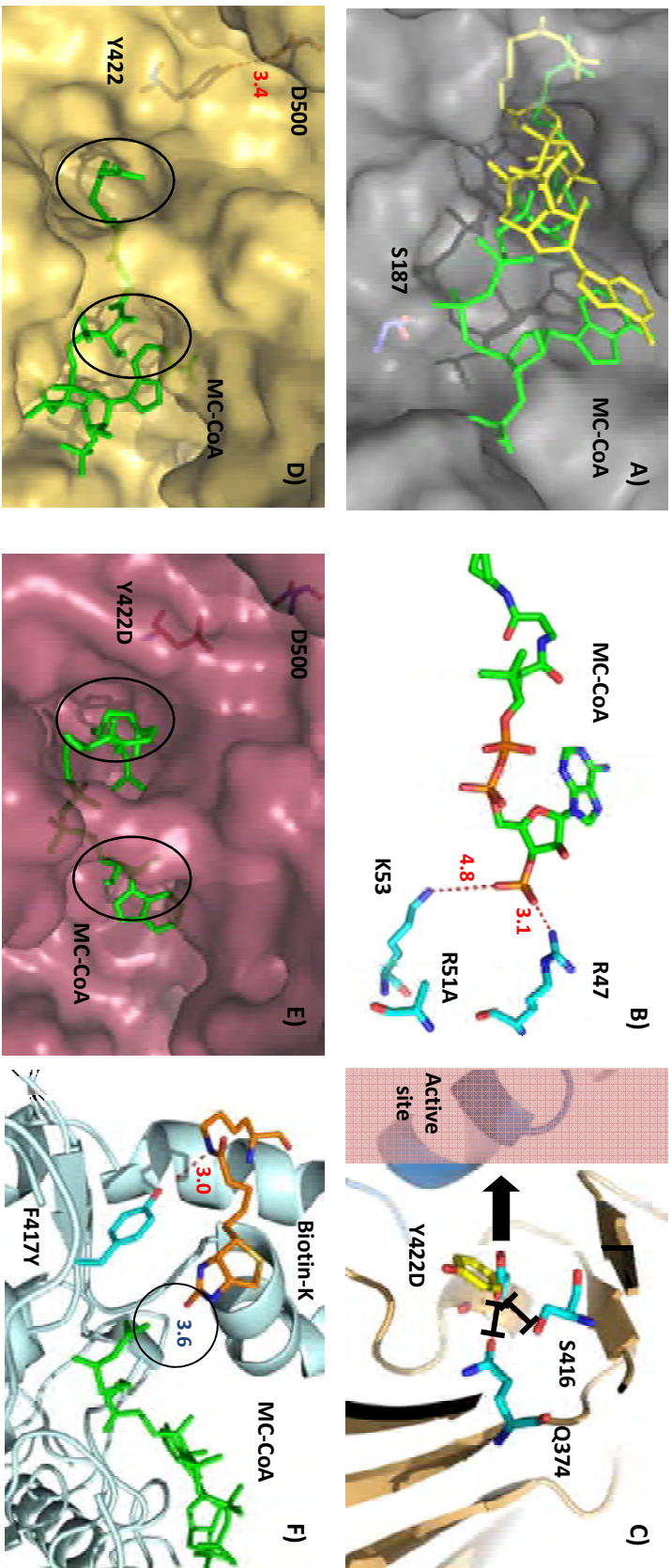


Figure 6. - Docking analysis of the mutants of MCCase of *P. aeruginosa*. A) Conformation of MC-CoA in the wild type MCCase (green) and in the mutant S187A (yellow). Docking analysis shows that the substrate not adopts an adequate conformation for the carboxylation reaction. B) Model of mutant R51A that indicates the probable interaction of MC-CoA substrate with residues K53 and R47, stabilizing the substrate. C) Replacement of Y422 by D can induce a repulsion effect of negative dipoles, and disrupts the active site. Black lines indicate the distance of repulsion, arrow indicates the direction of the repulsion. D) Interaction of residues Y422 and D500 in zone II of the active site and their coupling with the MC-CoA substrate. Pocket conformations are indicated with circles. E) Same as D), but with the Y422D mutation. F) Mutation F417Y in the active site of MCCase promotes the binding of biotin-lisine cofactor (orange) to the β' monomer located in zone II of the active site, where the transference of the carboxyl group to MC-CoA occurs. Distance between biotin and MC-CoA is indicated with a circle. The molecule of MC-CoA is colored green.

the role of Y422 in the MCCase is different to that found for the members of the BDC family, where it has been showed that it is involved in substrate recognition as reported for PCCase (Diacovich *et al.*, 2004). The FG double mutant showed V_{max} and K_m values 50% higher than the WT enzyme (Table 2), presenting an overlapping of effects as it was shown for the corresponding single mutants. That result suggests that the effect of activation of the mutation F417Y is counteracted by the effect of mutation G423A in the FG double mutant, since that V_{max} value of F417Y was the double compared with mutant FG (Table 2).

As mentioned above, the GCCase is an enzyme capable to use MC-CoA as alternative substrate (Aguilar *et al.*, 2008), therefore it was predicted that the mutations in zone II of the MCCase, changing the residues that resulted different after the comparison of the active site of the two enzymes, would cause that the MCCase mutant acquired GCCase activity. However, the results showed that the MCCase mutants designed were not capable to use G-CoA as substrate (data not shown), indicating that other residues than F417 and G423 are responsible for substrate discrimination in the *P. aeruginosa* MCCase.

In addition, recent oligomerization state studies showed that the enzyme MCCase lose the activity due to a change in the oligomerization state, forming aggregates of high molecular weight (Diaz-Perez *et al.*, in press). Results indicated that the lost of the activity in the S187A and Y422D MCCase mutants was not caused by a change in oligomerization state (data not shown). Recently, the crystal structure of the MCCase from *P. aeruginosa* with coenzyme A has been resolved (Hang *et al.*, 2011). Surprisingly in the crystal, the oligomeric state found was $\alpha_6\beta_6$, in contrast with the results obtained by gel filtration chromatography $\alpha_4\beta_4$ (Diaz-Perez *et al.*, in press). It is possible that the oligomerization found for the crystal of *P. aeruginosa* MCCase was formed by conversion of the active form $\alpha_4\beta_4$, to $\alpha_6\beta_6$ inactive form. Nevertheless, the active site localization and orientation in the model proposed in this work is the same than the resulting from the MCCase X-ray diffraction, suggesting that the dimeric functional unit of the MCCase β subunits is conserved independently of the oligomeric state.

Conclusions

In conclusion, a reliable tridimensional model of the MCCase β -subunit was obtained, showing that it was constituted for two closely homologous domains. This suggests that the β -subunit evolved from a common origin by a domain duplication event. The residues of the active site of the MCCase were identified by conducting structure and sequence alignments, site-directed mutagenesis and kinetic analyses. It was determined that the rol of residues R51 and S187of the MCCase is to bind and positioning the substrate MC-CoA in the active site, while that of residues F417, Y422 and G423 is involved in the formation of the acyl-biotin binding pocket. Furthermore, residues S187 and Y422 are essential for enzyme catalysis. The results reported in this work help to determine the importance of the residues located in the MCCase active site for the catalysis, in comparison to other BDC family members.

ACKNOWLEDGMENTS

We thank the *Pseudomonas* Genome Project and the *Pseudomonas aeruginosa* Community Annotation Project. This research was funded by CONACYT (83192, 106567) and C.I.C. 2.14/UMSNH grants.

REFERENCES

- Aguilar, J. A., Zavala, A. N., Diaz-Perez, C., Cervantes, C., Diaz-Perez, A. L. & Campos-Garcia, J. (2006). The *atu* and *liu* clusters are involved in the catabolic pathways for acyclic monoterpenes and leucine in *Pseudomonas aeruginosa*. *Appl Environ Microbiol* 72, 2070-2079.
- Aguilar, J. A., Diaz-Perez, C., Diaz-Perez, A. L., Rodriguez-Zavala, J. S., Nikolau, B. J. & Campos-Garcia, J. (2008). Substrate specificity of the 3-methylcrotonyl coenzyme A (CoA) and geranyl-CoA carboxylases from *Pseudomonas aeruginosa*. *J Bacteriol* 190, 4888-4893.

Campos-García, J. (2010). Metabolism of Acyclic Terpenes by *Pseudomonas*. In *Pseudomonas*, pp. 235-253. Edited by J. L. Ramos & A. Filloux: Springer Netherlands.

Diacovich, L., Mitchell, D. L., Pham, H., Gago, G., Melgar, M. M., Khosla, C., Gramajo, H. & Tsai, S. C. (2004). Crystal structure of the beta-subunit of acyl-CoA carboxylase: structure-based engineering of substrate specificity. *Biochemistry* 43, 14027-14036.

Díaz-Pérez, A. L., Zavala-Hernández, A. N., Cervantes, C. & Campos-García, J. (2004). The *gnyRDBHAL* cluster is involved in acyclic isoprenoid degradation in *Pseudomonas aeruginosa*. *Appl Environ Microbiol* 70, 5102-5110.

Díaz-Pérez, C., Rodríguez-Zavala, J. S., Díaz-Pérez, A. L. & Campos-García, J. (2011). Co-expression of α and β subunits of the 3-methylcrotonyl-coenzyme A carboxylase from *Pseudomonas aeruginosa*. *World J. Microbiol. Biotechnol.* Online First, 19 October 2011, DOI 10.1007/s11274-011-0921-1.

Fiser, A. & Sali, A. (2003). Modeller: generation and refinement of homology-based protein structure models. *Methods Enzymol* 374, 461-491.

Hall, P. R., Wang, Y. F., Rivera-Hainaj, R. E., Zheng, X., Pustai-Carey, M., Carey, P. R. & Yee, V. C. (2003). Transcarboxylase 12S crystal structure: hexamer assembly and substrate binding to a multienzyme core. *EMBO J* 22, 2334-2347.

Holm, L. & Rosenstrom, P. (2010). Dali server: conservation mapping in 3D. *Nucleic Acids Res* 38, W545-549.

Huang, C. S., Sadre-Bazzaz, K., Shen, Y., Deng, B., Zhou, Z. H. & Tong, L. (2010). Crystal structure of the $\alpha(6)\beta(6)$ holoenzyme of propionyl-coenzyme A carboxylase. *Nature* 466, 1001-1005.

Huang, C. S., Ge P., Zhou, Z. H. & Tong, L. (2011). An unanticipated architecture of the 750-kDa $\alpha\beta_6$ holoenzyme of 3-methylcrotonyl-CoA carboxylase. *Nature* (2011). doi:10.1038/nature10691.



Jitrapakdee, S. & Wallace, J. C. (2003). The biotin enzyme family: conserved structural motifs and domain rearrangements. *Curr Protein Pept Sci* 4, 217-229.

Laskowski, R. A., J. A. Rullmann, M. W. MacArthur, R. Kaptein & J. M. Thornton. (1996). AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J Biomol NMR*. 8(4), 477-486.

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

Luthy, R., Bowie, J. U. & Eisenberg, D. (1992). Assessment of protein models with three-dimensional profiles. *Nature* 356, 83-85.

Maiti, R., Van Domselaar, G. H., Zhang, H. & Wishart, D. S. (2004). SuperPose: a simple server for sophisticated structural superposition. *Nucleic Acids Res* 32, W590-594.

Morris, G. M., R. Huey, W. Lindstrom, M. F. Sanner, R. K. Belew, D. S. Goodsell & A. J. Olson. (2009). AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J Comput Chem*. 30(16), 2785-2791.

Notredame, C., Higgins, D. G. & Heringa, J. (2000). T-Coffee: A novel method for fast and accurate multiple sequence alignment. *J Mol Biol* 302, 205-217.

Sanner, M. F. (1999). Python: a programming language for software integration and development. *J Mol Graph Model*. 17(1), 57-61.

Toh, H., H. Kondo & T. Tanabe. (1993). Molecular evolution of biotin-dependent carboxylases. *Eur J Biochem*. 215(3), 687-696.

Trott, O. & A. J. Olson. (2010). AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J Comput Chem*. 31(2), 455-461.

Voss, N. R. & Gerstein, M. (2010). 3V: cavity, channel and cleft volume calculator and extractor. *Nucleic Acids Res* 38, W555-562.

Wiederstein, M. & Sippl, M. J. (2007). ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Res* 35, W407-410.

Zhang, H., Yang, Z., Shen, Y. & Tong, L. (2003). Crystal structure of the carboxyltransferase domain of acetyl-coenzyme A carboxylase. *Science* 299, 2064-2067.

Tables

Table 1.- Quality of structures for model and templates checked by PROCHECK.

		Ramachandran plot quality (%)			
		Core	Allowed	General	Disallowed
Model	LiuB	90.4	8.5	1.1	0.0
Template 1	1ON3	89.0	10.4	0.6	0.0
Template 2	1VRG	90.9	8.6	0.0	0.5
Template 3	1XNY	90.1	9.2	0.5	0.2

Ramachandran plot qualities show percentage (%) of residues belonging to the core, allowed, generally allowed and disallowed region of the plot.

Table 2.- Kinetics parameters of the *P. aeruginosa* MCCase recombinant enzyme and site-direct mutants.

	Km*	Vmax**	Vmax/Km***
WT	216 ± 30	477 ± 26	2.21
S187A	ND	ND	ND
R51A	309 ± 66	218 ± 20	0.71
F417Y	514 ± 80	1354 ± 296	2.64
FG (F417Y & G423A)	315 ± 105	711 ± 102	2.25
Y422D	ND	ND	ND
G423A	179 ± 83	327 ± 56	1.82

* μM , ** $\text{nmol CO}_2 \text{ fixed} \cdot \text{mg}^{-1} \cdot \text{min}^{-1}$, *** min^{-1} , ND: mutant where enzymatic activity was not shown.

SUPPLEMENTARY MATERIAL

Table S1.- Strain and plasmids used in this work.

Strain/Plamid	Characteristics	Reference
<i>Escherichia coli</i> JM101	Plasmid recipient strain	(Yanisch-Perron et al. 1985)
<i>E. coli</i> BL21(DE3)pLysS	Expression strain	(Studier y Moffatt 1986)
PCDB3	Co-expression plasmid of WT MCCase	(Díaz-Pérez et al. 2012)
PCYC-B	Plasmid pCYC184 containing fragment of <i>liuB</i> gene for mutagenesis	This work
PCDB-R51	Co-expression plasmid of R51A mutant MCCase	This work
PCDB-S187	Co-expression plasmid of S187A mutant MCCase	This work
PCDB-F417	Co-expression plasmid of F417Y mutant MCCase	This work
PCDB-Y422	Co-expression plasmid of Y422D mutant MCCase	This work
PCDB-G423	Co-expression plasmid of G423A mutant MCCase	This work
PCDB-FG	Co-expression plasmid of F417Y-G423A double mutant MCCase	This work

Table S2.- Oligonucleotides used in this work.

liuB-R51d	5' gcccgccacagcgcg gcc ggcaagctgctg 3'
liuB-R51r	5' cgcgctgtgccggcctgggcggc 3'
liuB-s187Ad	5' gtggtcatgggc gccc tcaccgccggc 3'
liuB-s187Ar	5' gcccatgaccacggcgcgctgctg 3'
liuB-F417Yd	5' ctgatcggcggcagc tac ggcgccggcaac 3'
liuB-F417Yr	5' gctgcccgatcagcagcggtgaa 3'
liuB-Y422Dd	5' ttcggcggcgaac gac ggcatgtgcgga 3'
liuB-Y422Dr	5' gttgccggcggcgaagctgccc 3'
liuB-G423Ad	5' ggccggcgaactac gcc atgtgcggacgc 3'
liuB-G423Ar	5' gtagtgccggcggcgaagctgcc 3'

Table 3.- Docking energies of MC-CoA to MCCase wt and mutants

	(kcal/mol)
WT	-8.2
R51	-7.2
S187	-7.3
F417	-7.7
Y422	-7.3
G423	-8.0
FG	-7.1

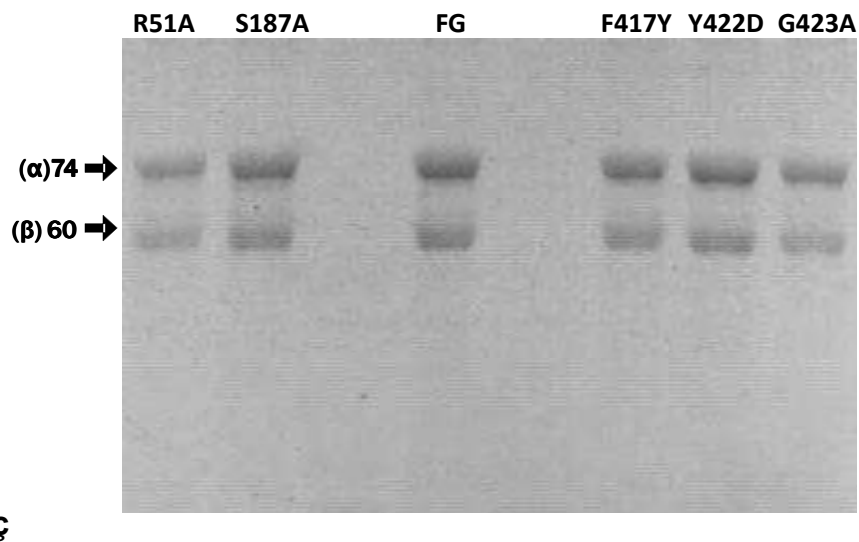


Figure S1.- Purification of the co-expressed and co-purified MCCase enzyme from *P. aeruginosa*. Coomassie blue-stained SDS-PAGE analysis of the purified recombinant α and β subunits of the MCCase and its mutant proteins are indicated.

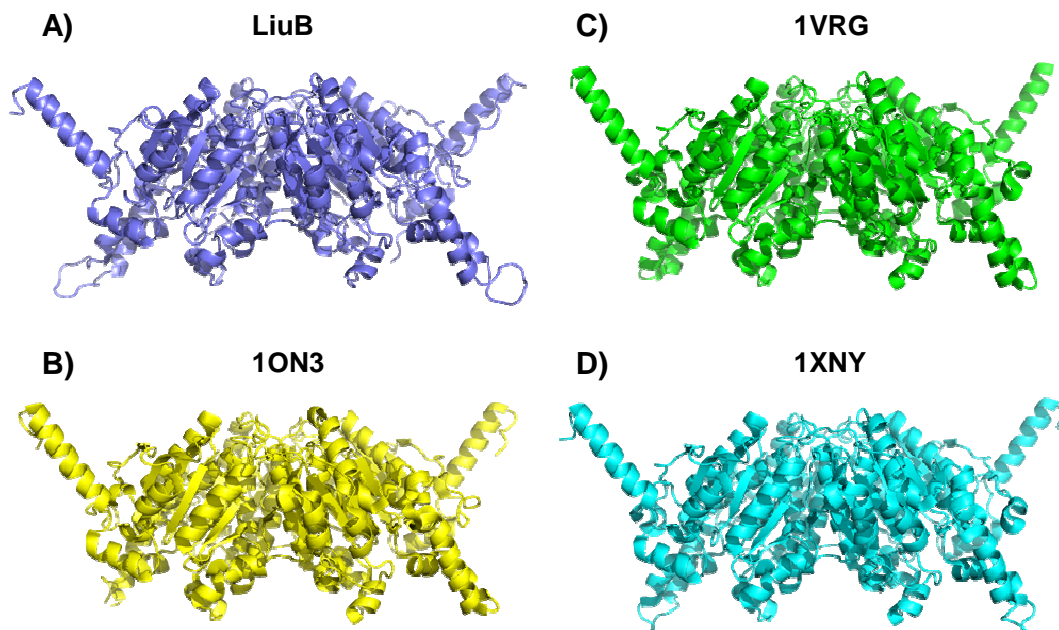


Figure S2.- Structural representation of the MCCase. A) *P. aeruginosa* molecular model (blue, PDB id - 1DQU, B) *P. freudenreichii* (yellow, 1F61); C) *T. maritime* (green, 1VRG); D) *S. coelicolor* (cyan - 1XNY).

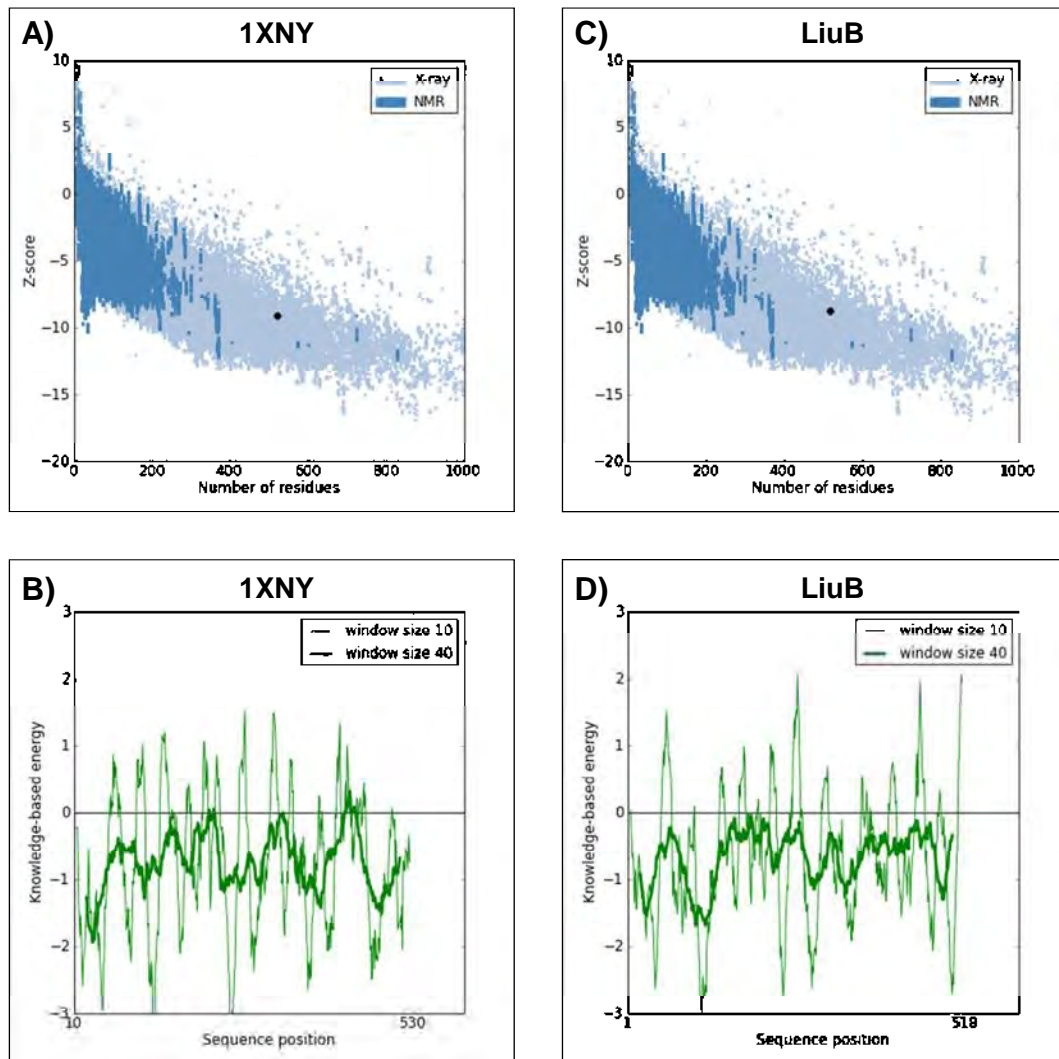


Figure S4.- ProSA energy profiles calculated for the template 1XNY (A-B) and the MODELLER model of β -subunit of MCCase (C-D).

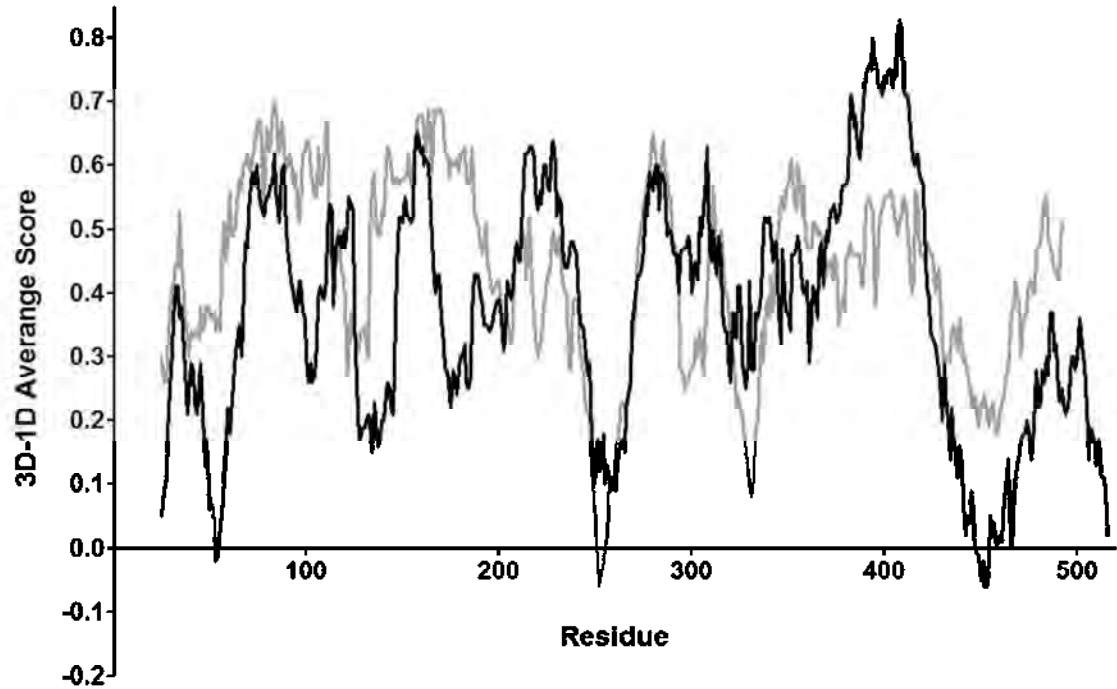


Figure S5.- Verified-3D analysis of template 1XNY (gray line) and MODELLER model of β -subunit of MCCase (black line).

VII.3.b) Métodos Suplementarios

VII.3.b.i) Modelaje iterativo

Se modeló la MCCasa de *P. aeruginosa* usando los tres moldes descritos anteriormente con el programa Modeller 9v6 (Fiser y Sali 2003), generando 10 modelos. Se eligió el mejor modelo obtenido, éste se agregó como molde para la segunda iteración y se corrigió el alineamiento de los moldes y la secuencia de la MCCasa. Se efectuó la segunda iteración generando otros 10 modelos. Los modelos obtenidos no pasaron la validación, por lo que se procedió a realizar un refinamiento a mano del alineamiento usado para generar los modelos, en las zonas donde se observaron los errores de validación. En este punto se incluyó dentro del alineamiento el mejor modelo no validado, al cual se le eliminaron las zonas erróneas, permitiendo al programa de modelado trabajar exhaustivamente solo en las zonas a corregir. Además, también se corrigieron los parámetros impuestos en la restricciones tridimensionales y en los parámetros de refinado impuestos por el programa de modelado, y se procedió a la generación de otros 10 modelos. Se realizaron cinco iteraciones para obtener un modelo que pasó la validación (Fig. 1).

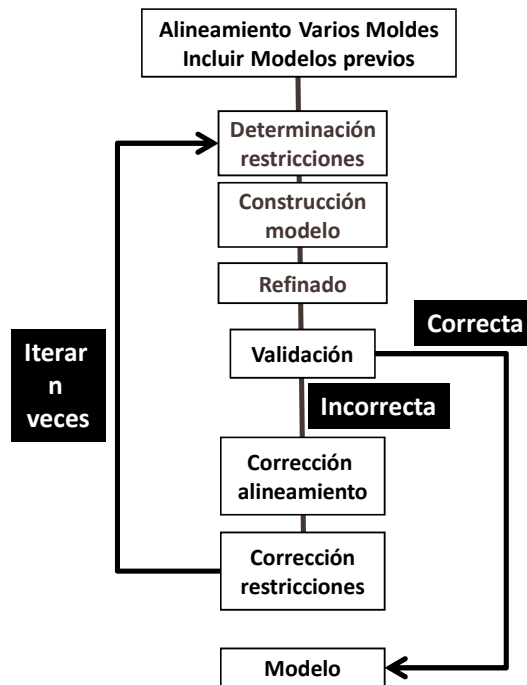


Figura 1.- Diagrama de flujo para la obtención de los modelos de la subunidad β de la MCCasa de *P. aeruginosa*.

VII.3.c) REFERENCIAS

- Díaz-Pérez, C., J. S. Rodríguez-Zavala, A. L. Díaz-Pérez y J. Campos-García. 2012. **Co-expression of α and β subunits of the 3-methylcrotonyl-coenzyme A carboxylase from *Pseudomonas aeruginosa*** World J. Microbiol. Biotechnol.
- Fiser, A. y A. Sali. 2003. **Modeller: generation and refinement of homology-based protein structure models.** Methods Enzymol. 374: 461-491.
- Studier, F. W. y B. A. Moffatt. 1986. **Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes.** J Mol Biol. 189(1): 113-130.
- Yanisch-Perron, C., J. Vieira y J. Messing. 1985. **Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors.** Gene. 33(1): 103-119.

IX) DISCUSIÓN GENERAL

En este trabajo se propuso la historia evolutiva de la subunidad β de la MCCasa de *P. aeruginosa*, la caracterización bioquímica de la enzima co-expresada, la localización y el estudio del papel de algunos de los aminoácidos que conforman el sitio activo del dominio de transcarboxilación.

IX.1) Historia evolutiva de las transcarboxilasas dependientes de coenzima-A (TDC)

Se encontró que las TDC son una superfamilia ampliamente distribuida en los tres dominios de la vida, localizada en 38 reinos distintos y 2556 miembros no redundantes, los cuales se distribuyen en dos linajes principales y 9 familias.

Las TDC se encargan del reconocimiento de substratos con CoA, es por esto que se consideró importante abordar la evolución de esta superfamilia tratando de encontrar una correlación de su distribución filogenética de acuerdo al substrato utilizado.

Se encontró que la superfamilia está compuesta de familias parálogas, claramente distribuidas de acuerdo al tipo químico de substrato utilizado, más aún, al inspeccionar a detalle el subgrupo donde se localiza la MCCasa de *P. aeruginosa*, se observó que se agrupa de manera independiente a la GCCasa de este organismo, por lo que es posible que la GCCasa haya surgido a partir de un evento de duplicación de la MCCasa, de tal forma la GCCasa se ha especializado en reconocer al G-CoA y ha conservado la capacidad de usar MC-CoA (Aguilar *et al.* 2008).

Las TDC son parte de un complejo multiproteico, en la MCCasa y GCCasa, el complejo es formado por la subunidad α , que lleva a cabo la fijación del CO₂, y la subunidad β , que contiene al dominio CT. En un estudio anterior se realizó la filogenia de la subunidad α (Aguilar *et al.* 2006), en donde se observó que la MCCasa y la GCCasa forman parte de un mismo grupo filogenético (Figura 1A). Al comparar la filogenia obtenida del subgrupo que contiene a las subunidades β de la MCCasa y GCCasa (Figura 1B), se observó que la distancia evolutiva de las subunidades β es

mayor a la distancia de las subunidades α , lo que sugiere una fuerza divergente mayor en las subunidades β . Los genes de ambas carboxilasas se encuentran formando operones, y en ambas enzimas hay una estrecha interacción física entre la subunidad α y β (Díaz-Pérez *et al.* 2004; Aguilar *et al.* 2006; Aguilar *et al.* 2008), por lo que se esperaría que ambas subunidades coevolucionaran formando un patrón de árboles en espejo (Yeang y Haussler 2007; Pazos y Valencia 2008), lo cual no ocurre. Una posible explicación a este fenómeno, es que la presión de selección hacia el reconocimiento de los sustratos sea diferente en ambas subunidades, ya que en la MCCasa y GCCasa las subunidades α reconocen los mismos compuestos, mientras en las subunidades β , el sustrato principal es diferente.

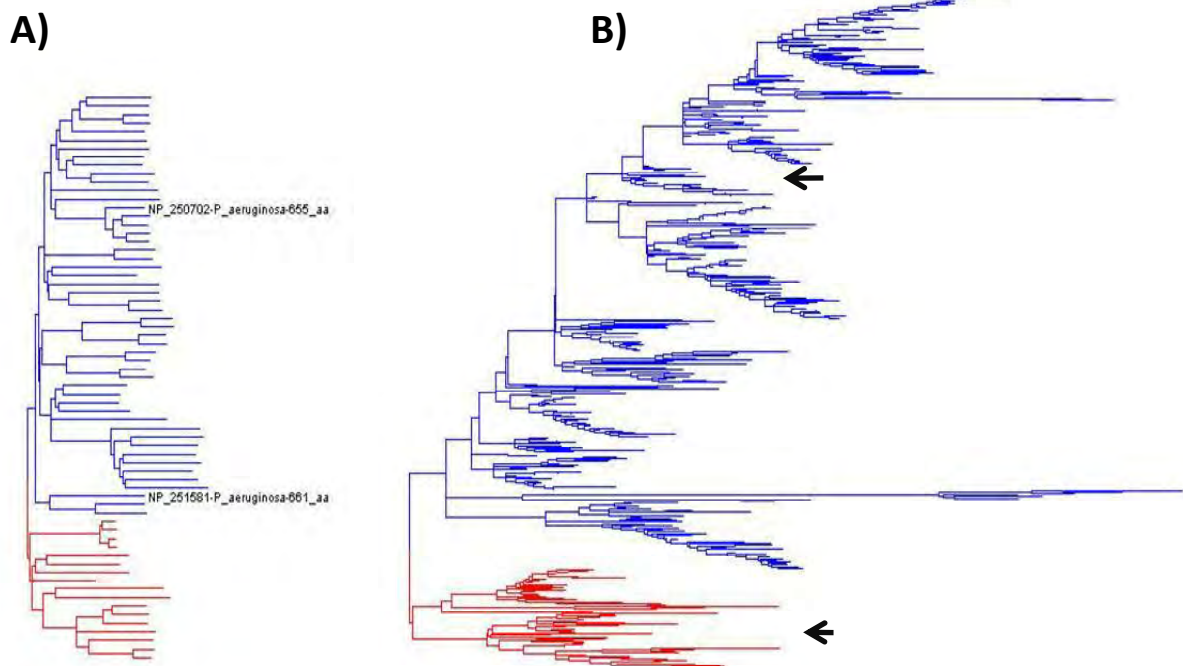


Figura 1.- Árboles filogenéticos de las subunidades α y β de las CDB. A) Árbol filogenético del grupo que contiene a la subunidad α de la MCCasa y GCCasa de *P. aeruginosa*. b) Árbol filogenético del subgrupo que contiene a la subunidad β de la MCCasa y GCCasa de *P. aeruginosa*. Se presenta con distinto color los subgrupos que forman al árbol. Las subunidades de la MCCasa (azul) y GCCasa (rojo) de *P. aeruginosa* se resaltan con fechas.

IX.2) Caracterización bioquímica de la MCCasa de *P. aeruginosa*

Para estudiar las características de cualquier proteína, es deseable obtenerla en su forma activa fácilmente, por esto se planteó la co-expresión de la MCCasa a partir de un solo vector que contiene los dos genes que la codifican, lográndose así obtener la enzima pura y activa eficientemente.

Es interesante que la MCCasa co-expresada presentó un comportamiento hiperbólico, en comparación con la MCCasa reconstituida, la cual mostró un comportamiento sigmoidal. Se observó que cuando la MCCasa no se incubaba previamente con el MC-CoA, existe un retraso en los primeros puntos en la gráfica de tiempo de reacción contra actividad enzimática. Esta variación de comportamiento enzimático puede ser un indicio de histéresis (Nelson-Rossow *et al.* 1993), la cual sería generada por la ausencia de un acoplamiento total de las subunidades, provocando un retraso en la actividad en bajas concentraciones de sustrato, por lo que al graficar la actividad contra la concentración de sustrato, se genera el comportamiento sigmoide. En las enzimas que presentan histéresis por falta de sustrato, este fenómeno es corregido cuando se incubaba a la enzima con el sustrato previo al inicio de la reacción, para propiciar el correcto acoplamiento de la enzima, como se hizo en este trabajo (Neet y Ainslie 1980).

La K_m y V_{max} obtenidas fueron de 168 μM para el MC-CoA y de 430 $nmol \cdot mg^{-1} \cdot min^{-1}$, valores que son parecidos a los de *P. citronellolis* (K_m de 138 μM), la MCCasa caracterizada de secuencia más parecida a la enzima de *P. aeruginosa*.

IX.3) Análisis del estado de oligomerización

La MCCasa de *P. aeruginosa* presentó una estructura cuaternaria $\alpha_4\beta_4$, al igual que la enzima de *P. citronellolis* (Fall 1981). La formación de un hetero-octámero es una forma intermedia entre las ACCasas ($\alpha_2\beta_2$) de *S. cerevisiae* y GDCasa ($\alpha_2\beta_2$) de *A. fermentans*, y las PCCasas ($\alpha_6\beta_6$), *S. coelicolor* (Zhang *et al.* 2003; Diacovich *et al.* 2004; Huang *et al.* 2010). Recientemente se caracterizó el cristal de la MCCasa de *P. aeruginosa*, en el cual la estructura observada es un dodecámero ($\alpha_6\beta_6$). La MCCasa

es muy inestable y se observa la formación de agregados de alto peso molecular después de 5 días de almacenamiento, lo cual puede explicar el por qué la enzima recién purificada y activa forme un octámero, detectado mediante filtración en gel, mientras que en la proteína cristalizada se forma un dodecámero. Sin embargo, no se descarta que exista un estado de interconversión entre los distintos estados de oligomerización en esta enzima.

IX.4) Modelaje de la MCCasa de *P. aeruginosa*

Si bien el estado de oligomerización es variable en la familia TDC, la unidad funcional de la subunidad β es un dímero, el cual es importante para la actividad y estabilización de los dominios CT (Hamed *et al.* 2008). El dímero de la subunidad β , además, es necesario para la formación del sitio activo localizado en un hueco creado en la zona de interacción entre las dos subunidades β y, debido a que la dimerización ocurre de manera cabeza-cola, se forman dos sitios activos por cada dímero de subunidades β (Hall *et al.* 2003; Diacovich *et al.* 2004). Más aún, en las proteínas cristalizadas como dímeros, y en las cristalizadas como hexámeros, el sitio activo es muy similar.

Para conocer el sitio activo de la MCCasa de *P. aeruginosa*, se llevó a cabo el modelaje molecular de esta proteína. El abordaje utilizado para la construcción del modelo fue implementando el uso combinado de varios moldes, un proceso iterativo de corrección de los alineamientos, verificación y uso de los mejores modelos obtenidos en la iteración siguiente, lo que llevó a la obtención de un modelo adecuado que pasó los estándares de tres programas diferentes de verificación.

Al igual que los moldes elegidos, el modelo generado de cada subunidad de la MCCasa presentó un plegamiento de crotonasa en cada uno de los subdominios que forman a la subunidad, con el plegamiento canónico de la superfamilia compuesto por repetidos de unidades $\beta\beta\alpha$ que se ensamblan en dos láminas β rodeadas por hélices α (Figura 2).

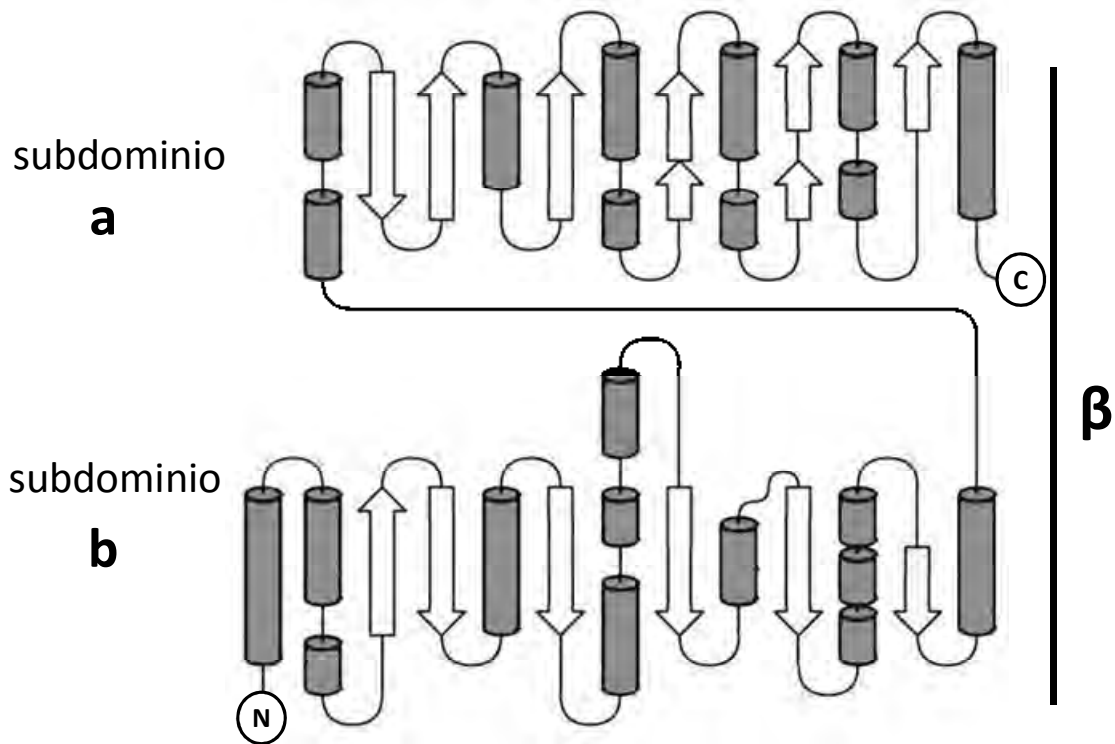


Figura 2.- Topología de la subunidad β de la MCCasa de *P. aeruginosa*. La subunidad está compuesta por dos dominios, a y b, con un plegamiento de crotonasa.

IX.5) Sitio activo de la MCCasa de *P. aeruginosa*

En la unidad funcional de las TDC, el dominio amino de una subunidad interacciona con el dominio carboxilo de la otra subunidad. En esta zona de interacción se ha localizado el sitio activo de transcarboxilación en la ACCasa, PCCasa y MTCasa (Hall *et al.* 2003; Zhang *et al.* 2003; Diacovich *et al.* 2004). Lo que nos sugirió que el sitio activo de la MCCasa se localiza también en esta zona.

Para localizar el sitio activo en el modelo generado de la MCCasa, se realizó un alineamiento estructural con el dímero de las subunidades β de la PCCasa de *S. coelicolor* (PDB:1XNY) y las subunidades 12S de la MTCasa de *P. freudenreichii* (PDB:1ON3). De esta forma se localizaron dos sitios activos en el modelo del dímero de la MCCasa, ambos sitios son prácticamente iguales, como ocurre en el cristal de

la proteínas moldes utilizadas, lo que sugiere que ambos sitios activos pueden ser utilizados simultáneamente para llevar a cabo la carboxilación del sustrato, por lo que la estequiometría de la enzima:sustrato es 1:1. Se observó que el sitio activo está formado por una gran cavidad (1494 Å³) en cada zona de interacción del dímero (Figura 3), siendo factible que una gran cantidad de residuos interactúen con el MC-CoA.

El análisis de anclaje molecular que se realizó arrojó observaciones interesantes. Primeramente el sitio activo puede ser dividido en dos regiones, la Zona I que une y dirige al MC-CoA mediante la interacción de los grupos fosfato de la CoA y un parche de aminoácidos básicos y polares de la enzima. Por otro lado, la zona II, donde se lleva a cabo la carboxilación de la cadena alifática del sustrato. En esta zona también interactúa la biotina, que lleva el grupo carboxilo que se transferirá al sustrato.

El análisis de anclaje molecular se acompañó con un alineamiento de la secuencia primaria de la MCCasa y varias TDC de función conocida. Ambos análisis nos guiaron a seleccionar a los residuos R51 y S187 en la zona I, debido a que son los que tienen una interacción más cercana a los grupos fosfato del MC-CoA y los residuos F417 y G423, dado que son variables en el sitio activo de la MCCasa en comparación de la GCCasa. Además, se seleccionó el residuo Y422, ya que se ha reportado que un cambio en esa posición afecta el sitio activo de la PCCasa de *S. coelicolor* (Diacovich *et al.* 2004).

IX.6) Papel de los residuos R51, S187, F417, Y422 y G423 en el sitio activo de la MCCasa de *P. aeruginosa*

Los resultados de la mutagénesis, modelado y anclaje molecular de las proteínas mutantes, en comparación con la proteína silvestre, arrojaron que la S187 es esencial para la funcionalidad de la MCCasa, ya que la mutación de este residuo genera una enzima inactiva. Se propone que S187 ayuda a que la molécula del sustrato se acomode para poder acoplarse al sitio activo. En cambio la mutación

R51A mostró actividad remanente, ya que en la zona donde se encuentra este residuo existen otros dos aminoácidos básicos que probablemente complementen parcialmente la función del residuo R51.

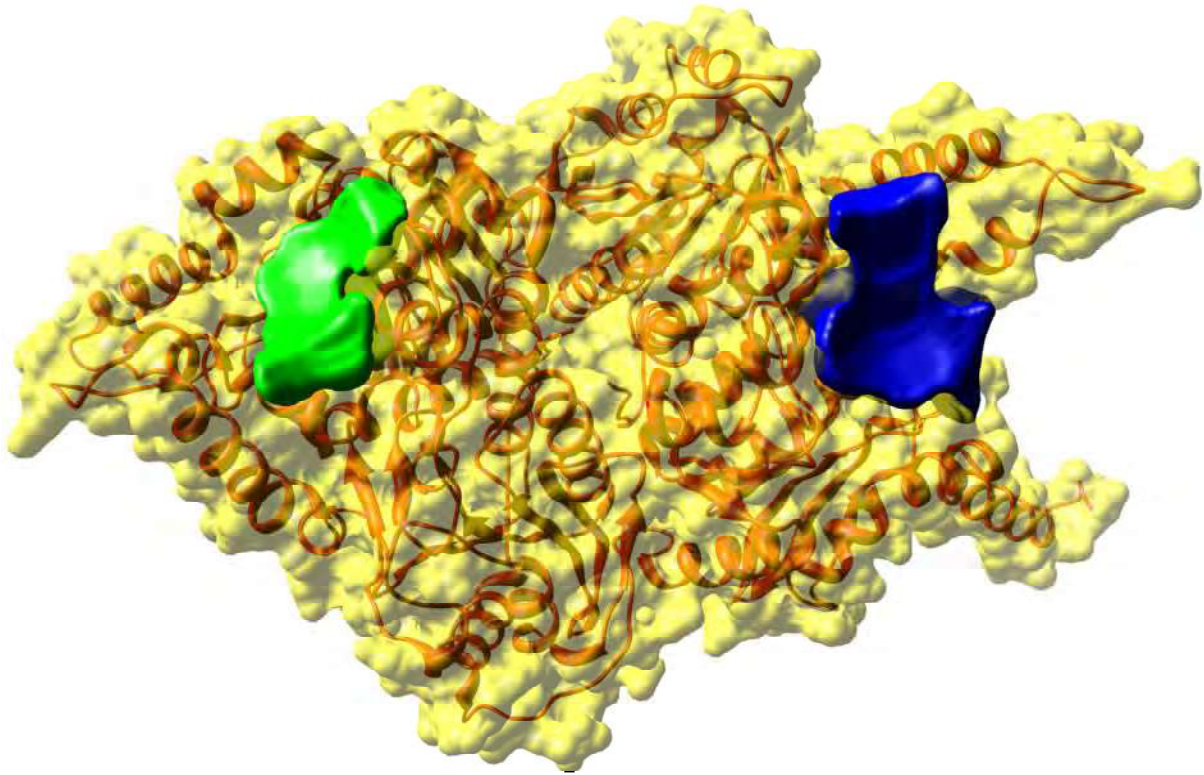


Figura 3.- Localización del sitio activo de la MCCasa de *P. aeruginosa*. El esqueleto de la MCCasa se muestra en naranja, y en amarillo se muestra la superficie expuesta al solvente. Las áreas donde se forman los dos sitios activos en cada dímero de subunidades β se presentan en verde y azul.

La mutación Y422D provoca que el sitio activo tenga una cavidad ligeramente mayor, ya que probablemente promueve el desplazamiento de un par de residuos lo que crea una pequeña protuberancia que impide el correcto posicionamiento del MC-CoA en el sitio activo, causando la inactivación completa de la enzima. Por otro lado, se propone que la mutación G423A alarga un poco una hélice- α en el sitio activo, provocando cambios leves en los parámetros cinéticos.

La mutación F417Y causa que aumente la V_{max} de la MCCasa, sin embargo, la K_m también aumenta. Este efecto parece ser debido a que la mutación favorece la unión de la biotina en el sitio activo, aumentando la velocidad de reacción, pero a su vez impide que la biotina salga del sitio activo y permanezca ocupado, lo que causa que la K_m aumente.

Las mutaciones simples F417Y y G423A, y la mutación doble FG (F417Y-G423A) no propiciaron el reconocimiento del G-CoA por la MCCasa, por lo que se concluye que otros residuos son los responsables de la discriminación del sustrato entre la MCCasa y GCCasa.

Estos resultados sugieren que los residuos R51, S187, F417, Y422 y G423 son parte importante del sitio activo de la MCCasa, y que es factible modificar los parámetros cinéticos de la enzima mediante la manipulación de estos residuos. Recientemente se resolvió el cristal de la MCCasa de *P. aeruginosa* (Huang et al. 2012), y se comparó la posición del sitio activo del modelo estructural generado en este trabajo con el del cristal, resultando ser muy similares. De la misma manera se comparó la posición del MC-CoA con el CoA presente en el cristal, lo que confirma los hallazgos en la cepas mutante de la zona I.

X) CONCLUSIONES

La superfamilia TDC es una familia ancestral, con dos linajes principales, los cuales se han diversificado para formar familias parálogas y posiblemente adaptarse al uso específico de compuestos acoplados a CoA. Los grupos de los linajes se han dividido de acuerdo al tipo de sustrato que reconocen. Así, la MCCasa de *P. aeruginosa* se encuentra en el linaje cCT agrupada con las GCCasas y GDCasas, la cuales tienen su acción en el carbono y del sustrato CoA.

Funcionalmente, al co-expresar las subunidades α y β de la MCCasa de *P. aeruginosa*, forma un complejo octamérico $\alpha_4\beta_4$, el cual presenta una cinética hiperbólica.

El modelo molecular de la subunidad β de la MCCasa de *P. aeruginosa* revela que los subdominios presentan el mismo plegamiento de crotonasa. El sitio activo se localiza en la interface de dos subunidades y se encuentra formado por dos zonas de unión, la primera encargada del posicionamiento y dirección del MC-CoA, y la segunda involucrada en la catálisis enzimática. En la primera zona la Serina 187 es indispensable para la funcionalidad de la enzima y la Arginina 51 es importante en la interacción con el MC-CoA. En la segunda zona la Tirosina 422 es indispensable para la función, mientras que la Fenilalanina 417 y la Glicina 423 ayudan en la interacción con el MC-CoA.

XI) PERSPECTIVAS

El presente trabajo nos permitió conocer la historia evolutiva de la subunidad β de la MCCasa de *P. aeruginosa*, así como el papel de algunos residuos del sitio activo de la misma. Sin embargo, como en todo trabajo científico, se generaron más preguntas, las cuales se resumen en este apartado.

XI.1) El complejo MCCasa

La MCCasa forma un complejo $\alpha_4\beta_4$, en comparación con las CDB cristalizadas cuyo estado de oligomerización es $\alpha_2\beta_2$ y $\alpha_6\beta_6$. Se sabe que las CDB que forman un complejo $\alpha_2\beta_2$ presentan la inserción de un dominio, el cual impide que se puedan formar complejos de mayor tamaño (Diacovich *et al.* 2004). Desde este punto de vista, sería relevante conocer cuáles son los determinantes estructurales que propician la formación del complejo $\alpha_4\beta_4$, dado que ya se cuenta con las condiciones para expresar y purificar la MCCasa rápida y fácilmente; además de que se encontró que el uso de trehalosa y de bajas concentraciones de imidazol permiten el almacenamiento de la proteína durante varios días sin la pérdida de la conformación y actividad, por lo que el reto sería encontrar las condiciones para la formación de cristales con la calidad necesaria para obtener la estructura de la proteína.

En este mismo contexto sería interesante caracterizar todas las superficies de contacto y localizar los aminoácidos responsables de la interacción tanto inter como intra subunidades α y β , así como la comparación con las CDB que formen complejos $\alpha_2\beta_2$ y $\alpha_6\beta_6$.

Debido a la complejidad del sistema, una alternativa que se ha tomado en muchas CDB es expresar y cristalizar las subunidades por separado, lo cual también nos daría información sobre la funcionalidad y los residuos importantes en ambas subunidades.

XI.2) La subunidad β de la MCCasa

Si bien se caracterizó el sitio activo de MCCasa, un planteamiento pendiente es averiguar cuál o cuáles son los aminoácidos responsables de que la MCCasa sea una enzima selectiva, mientras que su homóloga, la GCCasa, es una enzima promiscua. Para este fin se requiere de un análisis más profundo de la interacción entre la MCCasa y su substrato en la zona 2 de la subunidad β a partir de la estructura cristalizada de la MCCasa.

Otra duda que surgió, es conocer cuál es el papel de los demás aminoácidos que interaccionan con el MC-CoA en la zona I del sitio activo, además de la arginina 51. Dentro de estos residuos se encontraron a la arginina 47 y la lisina 53. Es factible que el MC-CoA interaccione con los 3 residuos, por lo que sería necesario generar mutantes sencillas y mutantes múltiples con combinaciones de estos residuos, para dilucidar su importancia en el sitio activo.

XI.3) La subunidad α de la MCCasa

Las subunidades α de las CDB presentan dos dominios claramente definidos y conservados, BC y BCCP. Además, en las PCCasas, MCCasas y GCCasa, se presenta un dominio de interconexión de ~100 aminoácidos. En el complejo PCCasa de *Ruegeria pomeroyi*, se ha encontrado que este dominio (denominado BT) sirve de puente para la interacción entre los dos sitios catalíticos localizados en los dominios BC y CT (Huang et al. 2010). La identidad en secuencia entre el dominio BT de las PCCasa y el dominio intermedio de la MCCasa es baja, por lo que no podemos saber si estos dominios tienen analogía funcional entre sí. Por lo que una de las perspectivas sería dilucidar el papel del dominio intermedio en la subunidad α de la MCCasa. Así mismo, se puede investigar si dicho dominio serviría como una huella genética que nos ayude a la discriminar a las carboxilasas sin función probada en las bases de datos.

XI.4) El complejo GCCasa

Dada la cercanía filogenética entre la GCCasa y la MCCasa, cabe suponer que la GCCasa también puede formar un complejo $\alpha_4\beta_4$, sin embargo, esto hay que demostrarlo experimentalmente. Para lo cual es necesario construir un plásmido de co-expresión que produzca eficientemente el complejo GCCasa.

Una propuesta que surge por la identidad de secuencias de las subunidades α de ambas carboxilasas, es probar si al ser intercambiadas estas subunidades, la función enzimática y/o la selectividad por el acil-CoA se mantiene, además de investigar su historia evolutiva completa.

XI.5) La familia de CDB

En el análisis filogenético de la subunidad β de las CDB, se encontraron homólogos de las PCCasas, GDCasas y MTCasas del dominio arquea. Hasta el momento, no se ha reportado la caracterización ni los productos génicos asociados a dichas enzimas, por lo que sería interesante llevar a cabo la caracterización de las distintas carboxilasas presentes en arqueas.

XII) REFERENCIAS

- Aguilar, J. A., C. Díaz-Pérez, A. L. Díaz-Pérez, J. S. Rodríguez-Zavala, B. J. Nikolau y J. Campos-García. 2008. **Substrate specificity of the 3-methylcrotonyl coenzyme A (CoA) and geranyl-CoA carboxylases from *Pseudomonas aeruginosa***. J Bacteriol. 190(14): 4888-4893.
- Aguilar, J. A., A. N. Zavala, C. Díaz-Pérez, C. Cervantes, A. L. Díaz-Pérez y J. Campos-García. 2006. **The *atu* and *liu* clusters are involved in the catabolic pathways for acyclic monoterpenes and leucine in *Pseudomonas aeruginosa***. Appl Environ Microbiol. 72(3): 2070-2079.
- Alban, C., P. Baldet, S. Axiotis y R. Douce. 1993. **Purification and characterization of 3-methylcrotonyl-Coenzyme A carboxylase from higher plant mitochondria**. Plant Physiology. 102(3): 957-965.
- Alexander, M. 1973. **Nonbiodegradable and other recalcitrant molecules**. Biotechnol. Bioeng. 15(4): 611 - 647.
- Campos-García, J. 2000. **Biodegradación y recalcitrancia de hidrocarburos**. Ciencia Nicolaita. 25: 79.
- Campos-García, J. 2010. **Metabolism of Acyclic Terpenes by *Pseudomonas***. En: J. L. Ramos and A. Filloux (Ed.) ***Pseudomonas***. Springer Netherlands. pp. 235-253.
- Cantwell, S. G., E. P. Lau, D. S. Watt y R. R. Fall. 1978. **Biodegradation of acyclic isoprenoids by *Pseudomonas* species**. J Bacteriol. 135(2): 324-333.
- Chapman-Smith, A., T. D. Mulhern, F. Whelan, J. E. Cronan, Jr. y J. C. Wallace. 2001. **The C-terminal domain of biotin protein ligase from *E. coli* is required for catalytic activity**. Protein Sci. 10(12): 2608-2617.
- Chavez-Aviles, M., A. L. Diaz-Perez y J. Campos-Garcia. 2010. **The bifunctional role of LiuE from *Pseudomonas aeruginosa*, displays additionally HIHG-CoA lyase enzymatic activity**. Mol Biol Rep. 37(4): 1787-1791.
- Chu, C. H. y D. Cheng. 2007. **Expression, purification, characterization of human 3-methylcrotonyl-CoA carboxylase (MCCC)**. Protein Expr Purif. 53(2): 421-427.
- Desviat, L. R., C. Perez-Cerda, B. Perez, J. Esparza-Gordillo, P. Rodriguez-Pombo, M. A. Penalva, S. Rodriguez De Cordoba y M. Ugarte. 2003. **Functional analysis of MCCA and MCCB mutations causing methylcrotonylglycinuria**. Mol Genet Metab. 80(3): 315-320.

- Diacovich, L., D. L. Mitchell, H. Pham, G. Gago, M. M. Melgar, C. Khosla, H. Gramajo y S. C. Tsai. 2004. **Crystal structure of the beta-subunit of acyl-CoA carboxylase: structure-based engineering of substrate specificity.** *Biochemistry*. 43(44): 14027-14036.
- Dias, F. F. y M. Alexander. 1971. **Effect of chemical structure on the biodegradability of aliphatic acids and alcohols.** *Appl Microbiol*. 22(6): 1114-1118.
- Díaz-Pérez, A. L. 2003. **Identificación de los genes de *Pseudomonas aeruginosa* involucrados en la biodegradación del isoprenoide citronelol.** Tesis de Maestría. IIQB. UMSNH. Morelia, Michoacán, México. 98 pp.
- Díaz-Pérez, A. L., A. N. Zavala-Hernández, C. Cervantes y J. Campos-García. 2004. **The gnyRDBHAL cluster is involved in acyclic isoprenoid degradation in *Pseudomonas aeruginosa*.** *Appl Environ Microbiol*. 70(9): 5102-5110.
- Fall, R. R. 1981. **3-Methylcrotonyl-CoA and geranyl-CoA carboxylases from *Pseudomonas citronellolis*.** *Methods Enzymol*. 71 Pt C: 791-799.
- Fall, R. R., J. L. Brown y T. L. Schaeffer. 1979. **Enzyme recruitment allows the biodegradation of recalcitrant branched hydrocarbons by *Pseudomonas citronellolis*.** *Appl Environ Microbiol*. 38(4): 715-722.
- Fall, R. R. y M. L. Hector. 1977. **Acyl-coenzyme A carboxylases. Homologous 3-methylcrotonyl-CoA and geranyl-CoA carboxylases from *Pseudomonas citronellolis*.** *Biochemistry*. 16(18): 4000-4005.
- Hall, P. R., Y. F. Wang, R. E. Rivera-Hainaj, X. Zheng, M. Pustai-Carey, P. R. Carey y V. C. Yee. 2003. **Transcarboxylase 12S crystal structure: hexamer assembly and substrate binding to a multienzyme core.** *EMBO J*. 22(10): 2334-2347.
- Hall, P. R., R. Zheng, L. Antony, M. Puzstai-Carey, P. R. Carey y V. C. Yee. 2004. **Transcarboxylase 5S structures: assembly and catalytic mechanism of a multienzyme complex subunit.** *EMBO J*. 23(18): 3621-3631.
- Hamed, R. B., E. T. Batchelar, I. J. Clifton y C. J. Schofield. 2008. **Mechanisms and structures of crotonase superfamily enzymes--how nature controls enolate and oxyanion reactivity.** *Cell Mol Life Sci*. 65(16): 2507-2527.
- Hoschle, B., V. Gnau y D. Jendrossek. 2005. **Methylcrotonyl-CoA and geranyl-CoA carboxylases are involved in leucine/isovalerate utilization (Liu) and acyclic terpene utilization (Atu), and are encoded by liuB/liuD and atuC/atuF, in *Pseudomonas aeruginosa*.** *Microbiology*. 151(Pt 11): 3649-3656.

- Huang, C. S., P. Ge, Z. H. Zhou y L. Tong. 2012. **An unanticipated architecture of the 750-kDa $\alpha_6\beta_6$ holoenzyme of 3-methylcrotonyl-CoA carboxylase.** Nature. 481(7380): 219-223.
- Huang, C. S., K. Sadre-Bazzaz, Y. Shen, B. Deng, Z. H. Zhou y L. Tong. 2010. **Crystal structure of the $\alpha_6\beta_6$ holoenzyme of propionyl-coenzyme A carboxylase.** Nature. 466(7309): 1001-1005.
- Jitrapakdee, S. y J. C. Wallace. 2003. **The biotin enzyme family: conserved structural motifs and domain rearrangements.** Curr Protein Pept Sci. 4(3): 217-229.
- Jordan, I. K., K. Henze, N. D. Fedorova, E. V. Koonin y M. Y. Galperin. 2003. **Phylogenomic analysis of the *Giardia intestinalis* transcarboxylase reveals multiple instances of domain fusion and fission in the evolution of biotin-dependent enzymes.** J Mol Microbiol Biotechnol. 5(3): 172-189.
- Lau, E. P., B. C. Cochran y R. R. Fall. 1980. **Isolation of 3-methylcrotonyl-coenzyme A carboxylase from bovine kidney.** Arch Biochem Biophys. 205(2): 352-359.
- Lau, E. P., B. C. Cochran, L. Munson y R. R. Fall. 1979. **Bovine kidney 3-methylcrotonyl-CoA and propionyl-CoA carboxylases: each enzyme contains nonidentical subunits.** Proc Natl Acad Sci U S A. 76(1): 214-218.
- Li, Z. G., W. B. Yin, H. Guo, L. Y. Song, Y. H. Chen, R. Z. Guan, J. Q. Wang, R. R. Wang y Z. M. Hu. 2010. **Genes encoding the alpha-carboxyltransferase subunit of acetyl-CoA carboxylase from *Brassica napus* and parental species: cloning, expression patterns, and evolution.** Genome. 53(5): 360-370.
- NCR (2003). **Oil in the Sea III: Inputs, Fates, and Effects.** Washington, DC. USA, The National Academies Press.
- Neet, K. E. y G. R. Ainslie, Jr. 1980. **Hysteretic enzymes.** Methods Enzymol. 64: 192-226.
- Nelson-Rossow, K. L., K. A. Sukalski y R. C. Nordlie. 1993. **Hysteresis at near-physiologic substrate concentrations underlies apparent sigmoid kinetics of the glucose-6-phosphatase system.** Biochim Biophys Acta. 1163(3): 297-302.
- Nikolau, B. J., J. B. Ohlrogge y E. S. Wurtele. 2003. **Plant biotin-containing carboxylases.** Arch Biochem Biophys. 414(2): 211-222.

- Pazos, F. y A. Valencia. 2008. **Protein co-evolution, co-adaptation and interactions**. EMBO J. 27(20): 2648-2655.
- Pirnik, M. P., R. M. Atlas y R. Bartha. 1974. **Hydrocarbon metabolism by *Brevibacterium erythrogenes*: normal and branched alkanes**. J Bacteriol. 119(3): 868-878.
- Rodriguez, J. M., P. Ruiz-Sala, M. Ugarte y M. A. Penalva. 2004. **Fungal metabolic model for 3-methylcrotonyl-CoA carboxylase deficiency**. J Biol Chem. 279(6): 4578-4587.
- Rojo, F. 2009. **Degradation of alkanes by bacteria**. Environ Microbiol. 11(10): 2477-2490.
- Rojo, F. 2010. **Carbon catabolite repression in *Pseudomonas* : optimizing metabolic versatility and interactions with the environment**. FEMS Microbiol Rev. 34(5): 658-684.
- Schaeffer, T. L., S. G. Cantwell, J. L. Brown, D. S. Watt y R. R. Fall. 1979. **Microbial growth on hydrocarbons: terminal branching inhibits biodegradation**. Appl Environ Microbiol. 38(4): 742-746.
- Schneiker, S., V. A. Martins dos Santos, D. Bartels, T. Bekel, M. Brecht, J. Buhrmester, T. N. Chernikova, R. Denaro, M. Ferrer, C. Gertler, A. Goesmann, O. V. Golyshina, F. Kaminski, A. N. Khachane, S. Lang, B. Linke, A. C. McHardy, F. Meyer, T. Nechitaylo, A. Puhler, D. Regenhardt, O. Rupp, J. S. Sabirova, W. Selbitschka, M. M. Yakimov, K. N. Timmis, F. J. Vorholter, S. Weidner, O. Kaiser y P. N. Golyshin. 2006. **Genome sequence of the ubiquitous hydrocarbon-degrading marine bacterium *Alcanivorax borkumensis***. Nat Biotechnol. 24(8): 997-1004.
- St Maurice, M., L. Reinhardt, K. H. Surinya, P. V. Attwood, J. C. Wallace, W. W. Cleland y I. Rayment. 2007. **Domain architecture of pyruvate carboxylase, a biotin-dependent multifunctional enzyme**. Science. 317(5841): 1076-1079.
- Stover, C. K., X. Q. Pham, A. L. Erwin, S. D. Mizoguchi, P. Warrener, M. J. Hickey, F. S. Brinkman, W. O. Hufnagle, D. J. Kowalik, M. Lagrou, R. L. Garber, L. Goltry, E. Tolentino, S. Westbrook-Wadman, Y. Yuan, L. L. Brody, S. N. Coulter, K. R. Folger, A. Kas, K. Larbig, R. Lim, K. Smith, D. Spencer, G. K. Wong, Z. Wu, I. T. Paulsen, J. Reizer, M. H. Saier, R. E. Hancock, S. Lory y M. V. Olson. 2000. **Complete genome sequence of *Pseudomonas aeruginosa* PAO1, an opportunistic pathogen**. Nature. 406(6799): 959-964.
- Toh, H., H. Kondo y T. Tanabe. 1993. **Molecular evolution of biotin-dependent carboxylases**. Eur J Biochem. 215(3): 687-696.



- Tu, Z. y H. H. Hagedorn. 1997. **Biochemical, molecular, and phylogenetic analysis of pyruvate carboxylase in the yellow fever mosquito, *Aedes aegypti***. *Insect Biochem Mol Biol.* 27(2): 133-147.
- Van Beilen, J. B., Z. Li, W. A. Duetz, T. H. M. Smits y B. Witholt. 2003. **Diversité des systèmes alcane hydroxylase dans l'environnement**. *Oil & Gas Science and Technology - Rev. IFP.* 58(4): 427-440.
- van Beilen, J. B., S. Panke, S. Lucchini, A. G. Franchini, M. Rothlisberger y B. Witholt. 2001. **Analysis of *Pseudomonas putida* alkane-degradation gene clusters and flanking insertion sequences: evolution and regulation of the alk genes**. *Microbiology.* 147(Pt 6): 1621-1630.
- Wentzel, A., T. E. Ellingsen, H. K. Kotlar, S. B. Zotchev y M. Throne-Holst. 2007. **Bacterial metabolism of long-chain n-alkanes**. *Appl Microbiol Biotechnol.* 76(6): 1209-1221.
- Wurtele, E. S. y B. J. Nikolau. 2000. **Characterization of 3-methylcrotonyl-CoA carboxylase from plants**. *Methods Enzymol.* 324: 280-292.
- Yakimov, M. M., P. N. Golyshin, S. Lang, E. R. Moore, W. R. Abraham, H. Lunsdorf y K. N. Timmis. 1998. ***Alcanivorax borkumensis* gen. nov., sp. nov., a new, hydrocarbon-degrading and surfactant-producing marine bacterium**. *Int J Syst Bacteriol.* 48 Pt 2: 339-348.
- Yeang, C. H. y D. Haussler. 2007. **Detecting coevolution in and among protein domains**. *PLoS Comput Biol.* 3(11): e211.
- Zhang, H., Z. Yang, Y. Shen y L. Tong. 2003. **Crystal structure of the carboxyltransferase domain of acetyl-coenzyme A carboxylase**. *Science.* 299(5615): 2064-2067.

Anexo I

Otros trabajos publicados durante la estancia en el PIDCB.

- Díaz-Pérez, A. L., C. Roman-Doval, **C. Díaz-Pérez**, C. Cervantes, C. R. Sosa-Aguirre, J. E. López-Meza y J. Campos-García. 2007. **Identification of the aceA gene encoding isocitrate lyase required for the growth of *Pseudomonas aeruginosa* on acetate, acyclic terpenes and leucine.** FEMS Microbiol Lett. 269(2): 309-316.
- Aguilar, J. A., **C. Díaz-Pérez**, A. L. Díaz-Pérez, J. S. Rodríguez-Zavala, B. J. Nikolau y J. Campos-García. 2008. **Substrate specificity of the 3-methylcrotonyl coenzyme A (CoA) and geranyl-CoA carboxylases from *Pseudomonas aeruginosa*.** Journal of Bacteriology. 190(14): 4888-4893.
- Ramírez-Díaz, M. I., **C. Díaz-Pérez**, E. Vargas, H. Riveros-Rosas, J. Campos-García y C. Cervantes. 2008. **Mechanisms of bacterial resistance to chromium compounds.** Biometals. 21(3): 321-332.
- Aguilar-Barajas, E., **C. Díaz-Pérez**, M. I. Ramírez-Díaz, H. Riveros-Rosas y C. Cervantes. 2011. **Bacterial transport of sulfate, molybdate, and related oxyanions.** Biometals. 24(4): 687-707.
- Ortiz-Castro, R., **C. Diaz-Perez**, M. Martinez-Trujillo, R. E. del Rio, J. Campos-Garcia y J. Lopez-Bucio. 2011. **Transkingdom signaling based on bacterial cyclodipeptides with auxin activity in plants.** Proc Natl Acad Sci U S A. 108(17): 7253-7258.