



**UNIVERSIDAD NACIONAL AUTÓNOMA DE MEXICO**

**PROGRAMA DE MAESTRÍA Y DOCTORADO EN CIENCIAS QUÍMICAS**

**CLONACIÓN, SOBREEXPRESIÓN Y CARACTERIZACIÓN CINÉTICA DE LA  
ENZIMA ALDEHÍDO DESHIDROGENASA PA2125 DE *Pseudomonas aeruginosa***

**TESIS**

PARA OPTAR POR EL GRADO DE

**MAESTRO EN CIENCIAS**

PRESENTA

**Químico CARLOS ALBERTO RUEDA SANABRIA**

Tutora: Doctora ROSARIO ADELAIDA MUÑOZ CLARES  
FACULTAD DE QUÍMICA, UNAM

CIUDAD DE MÉXICO, NOVIEMBRE DE 2017



Universidad Nacional  
Autónoma de México



**UNAM – Dirección General de Bibliotecas**  
**Tesis Digitales**  
**Restricciones de uso**

**DERECHOS RESERVADOS ©**  
**PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL**

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.



**UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO**

**PROGRAMA DE MAESTRÍA Y DOCTORADO EN CIENCIAS QUÍMICAS**

**CLONACIÓN, SOBREEXPRESIÓN Y CARACTERIZACIÓN CINÉTICA DE LA  
ENZIMA ALDEHÍDO DESHIDROGENASA PA2125 DE *Pseudomonas aeruginosa***

**T E S I S  
PARA OPTAR POR EL GRADO DE**

**MAESTRO EN CIENCIAS**

**P R E S E N T A**

**Químico CARLOS ALBERTO RUEDA SANABRIA**



## **JURADO**

**PRESIDENTE** Dra. Adela Rodríguez Romero

**VOCAL** Dr. Daniel Alejandro Fernández Velasco

**VOCAL** Dr. José Luis Medina Franco

**VOCAL** Dra. Karina Martínez Mayorga

**SECRETARIO** Dra. Nuria Victoria Sánchez Puig

Tutor principal

**Dra. Rosario A. Muñoz Clares**

\_\_\_\_\_

Sustentante

**Q. Carlos Alberto Rueda Sanabria**

\_\_\_\_\_

## AGRADECIMIENTOS

A la Universidad Nacional Autónoma de México (UNAM), a la Facultad de Química y al Programa de Maestría y Doctorado en Ciencias Químicas por la formación académica brindada y la posibilidad de poder llevar a cabo estudios en esta Alma Mater de tanto prestigio

Al Programa de Apoyo al Posgrado y la Innovación Tecnológica (PAPIIT) de la Dirección General de Apoyo al Personal Académico (DGAPA), UNAM, por el apoyo financiero otorgado en el proyecto IN220317.

Al Consejo Nacional de Ciencia y Tecnología (CONACyT) por la beca de posgrado otorgada bajo el número de becario 595323, así como por el apoyo económico otorgado a través del proyecto de infraestructura 212513.

A la Dra. Rosario Muñoz Clares, por permitirme hacer parte de su grupo de investigación, por su apoyo en los momentos difíciles, por sus enseñanzas y su acompañamiento día a día durante el desarrollo de este trabajo.

A los miembros del jurado, por las observaciones y sugerencias para mejorar la versión final del documento.

Al Biol. Carlos Mújica Jiménez, quien fue un excelente asesor técnico, un amigo y un compañero incondicional en el desarrollo de este trabajo.

A la Dra. Lilian González Segura, por sus enseñanzas en la determinación de estructuras cristalográficas de las enzimas y su asesoría en el manejo de programas para visualización de proteínas.

Al Dr. Rogelio Rodríguez Sotres, por la obtención de un modelo por homología de la estructura tridimensional de la enzima.

Al M. en C. Javier Carrillo Campos, por su ayuda y enseñanza en la parte de biología molecular de este trabajo.

A la Dra. Ana Juárez Vázquez, por sus enseñanzas en evolución de proteínas y todo el análisis bioinformático de la enzima *PA2125*.

A los compañeros del Laboratorio de Estructura, Función y Evolución de Proteínas (LEFEP), a Yudy Vanesa Cardona por su compañerismo, amistad y apoyo en los momentos difíciles, a Erick Palomo Paz, a Erik Fernando López, a Esdras Carrizosa, nuevamente al Msc. Javier Carrillo compañero incondicional, a Sandra Villegas, a Karla Casanova y a José Antonio Espinosa. A tod@s gracias por su ayuda y sus comentarios en la realización de este proyecto y, sobre todo, por los momentos compartidos y por hacerme sentir como un mexicano más.

.

## **DEDICATORIA**

*Dedico esta tesis a...*

*Mi hermosa esposa Mónica, mi gran amor quién con su apoyo y amor incondicional hizo que nunca decayera, e hizo lo imposible para que realizara mis estudios de maestría.*

*A mis hijos Santiago, Felipe y Nicolás quienes fueron mi motor y orgullo para culminar con esta meta y mostrarles que caminos fáciles hay muchos pero el más difícil es el más certero.*

*A mi madre Myriam y mi abuela Carmen por su apoyo, por sus oraciones, por su gran amor brindado sin esperar nada a cambio y por creer en mí.*

*A mi suegra Zoraida quién con su comprensión y apoyo estuvo con nosotros siempre.*

*A mis hermanos Adriana y Ricardo, a mi sobrina Aura María, a mis cuñados Miguel*

*Ángel, Héctor y Alejita gracias por su amor y confianza.*

## ÍNDICE

I ANTECEDENTES	19
I.1 Generalidades de las enzimas aldehído deshidrogenasas	19
I.2 Mecanismo químico de reacción catalizada por las aldehído deshidrogenasas	22
I.3 Mecanismo cinético	25
I.4 Propiedades estructurales de las aldehído deshidrogenasas	26
I.4.1 Estructura primaria	26
I.4.2 Plegamiento de los monómeros	26
I.4.3 Estructuras cuaternarias	32
I.5 Cinética del pre-estado estacionario de las aldehído deshidrogenasas	33
I.6 Género <i>Pseudomonas</i>	36
I.6.1 <i>Pseudomonas aeruginosa</i>	37
I.7 Enzimas aldehído deshidrogenasas de <i>Pseudomonas aeruginosa</i> PAO1	38
II JUSTIFICACIÓN DEL PROYECTO	39
III HIPÓTESIS	40
IV OBJETIVOS	41
IV.1 Objetivo general	41
IV.2 Objetivos particulares	41
V MATERIALES Y MÉTODOS	42
V.1 Estrategia experimental	42



V.2 Búsqueda de ortólogos y alineamientos múltiples de secuencias de aminoácidos	43
V.3 Clonación del gen <i>pa2125</i>	44
V.4 Sobreexpresión del gen <i>pa2125</i> y producción de la enzima recombinante <i>pa2125</i> en células de <i>E. coli</i>	49
V.5 Purificación a homogeneidad de la enzima <i>PA2125</i>	51
V.6 Determinación del estado oligomérico de la enzima <i>PA2125</i>	53
V.7 Ensayos de actividad y determinación de los parámetros cinéticos para diferentes sustratos aldehídos	54
V.8 Ensayos de cinética rápida	55
V.9 Obtención de un modelo por homología de la estructura tridimensional de la enzima <i>PA2125</i>	56
V.10 Acoplamiento molecular de los sustratos en el sitio activo de <i>PA2125</i>	57
VI RESULTADOS	59
VI.1 Ortólogos de <i>PA2125</i>	59
VI.2 Clonación del gen <i>PA2125</i>	72
VI.3 Sobreexpresión del gen <i>pa2125</i> y producción de la enzima recombinante <i>PA2125</i> en células de <i>E. coli</i>	75
VI.4 Purificación a homogeneidad de la enzima <i>PA2125</i>	77
VI.5 Estado oligomérico de la enzima <i>PA2125</i>	79
VI.6 Especificidad por el sustrato aldehído	80
VI.7 Parámetros cinéticos aparentes para benzaldehído y derivados	83

VI.8 Mecanismo cinético de <i>pa2125</i> usando benzaldehído como sustrato	87
VI.9 Cinética rápida	89
VI.10 Modelo por homología de la estructura tridimensional de la enzima <i>PA2125</i>	92
VI.11 Acoplamiento molecular de los posibles sustratos aldehídos en el sitio activo de la enzima <i>PA2125</i>	101
VII DISCUSIÓN DE RESULTADOS	108
VII.1 Papel fisiológico de la <i>PA2125</i>	108
VII.2 Mecanismo catalítico de la <i>PA2125</i>	115
VII.3 Mecanismo cinético de <i>PA2125</i>	121
VIII CONCLUSIONES	125
IX PERSPECTIVAS	127
X BIBLIOGRAFÍA	129
XI ANEXOS	144
Anexo 1. Genes que codifican para enzimas de la familia ALDH cd07120 contenidos en la base de datos del NCBI.	145
Anexo 2. Alineamientos de secuencias de aminoácidos entre <i>PA2125</i> , gentisaldehído deshidrogenasa ( <i>GTSALDH</i> ) y furfural deshidrogenasa ( <i>FURALDH</i> ).	156
Anexo 3. Alineamientos de secuencias de aminoácidos entre <i>PA2125</i> y enzimas que presentan actividad con aldehídos aromáticos derivados del <i>BZAL</i> , con succínico semialdehído, y <i>BAL</i> .	157

Anexo 4. Alineamiento de la secuencia de nucleótidos obtenida del gen clonado a partir del DNA genómico de <i>P. aeruginosa</i> con la secuencia del gen <i>pa2125</i> depositada en la base datos del NCBI.	159
Anexo 5. Resultados de la evaluación de estructuras cristalográficas como posibles moldes para la obtención del modelo por homología de <i>PA2125</i> .	161
Anexo 6. Puntuaciones RD.HMM de secuencias completas recuperadas por el modelo de Markov oculto de la base de datos de secuencias NR NCBI.	161
Anexo 7. Resultados de los experimentos de docking con aldehídos aromáticos derivados del BZAL.	162
Anexo 8. Comparación contextos genómicos de enzimas de la familia cd07120	166

## ÍNDICE DE FIGURAS

Figura 1. Estructura tridimensional de las enzimas aldehído deshidrogenasas.	27
Figura 2. Sitio activo de las aldehído deshidrogenasas mostrando los residuos catalíticos y las conformaciones posibles de la cisteína y el glutámico catalíticos.	30
Figura 3. Tipos de estructuras cuaternarias encontradas en aldehído deshidrogenasas.	32
Figura 4. Ejemplo de cursos temporales de enzimas aldehído deshidrogenasas en ensayos de cinética rápida.	35
Figura 5. Estrategia experimental planteada para el desarrollo de la investigación de este trabajo de tesis.	42
Figura 6 Mapa del vector pET-28b (+) en el que se clonó el gen <i>pa2125</i> .	47
Figura 7. Logos de los residuos conservados de la familia ALDH cd07120.	68
Figura 8. Contextos genómicos de los genes que codifican para <i>PA2125</i> , las enzimas de su misma familia con actividad probada y la <i>PaBADH</i> .	70
Figura 9. Gel de agarosa en el que se muestra la amplificación por PCR del gen <i>pa2125</i> .	73
Figura 10. Gel de agarosa mostrando los productos de la digestión del vector pET28b(+) y del gen <i>pa2125</i> usando las enzimas NdeI y HindIII.	73
Figura 11. Gel de agarosa mostrando la digestión del plásmido pET28- <i>pa2125</i> usando las enzimas NdeI y HindIII.	75
Figura 12. Gel SDS-PAGE mostrando la producción de la enzima <i>PA2125</i> recombinante en células de <i>E. coli</i> BL21 (DE3) Rosetta.	77
Figura 13. Gel SDS-PAGE mostrando el resultado de la purificación de la enzima <i>PA2125</i> .	78

Figura 14. Perfil de elución de la enzima <i>PA2125</i> de la cromatografía de exclusión molecular.	80
Figura 15. Actividad de la enzima <i>PA2125</i> con diversos aldehídos a concentración fisiológica de NAD <sup>+</sup> .	82
Figura 16. Cinéticas de saturación de la enzima <i>PA2125</i> con los aldehídos como sustrato variable.	84
Figura 17. Cinéticas de saturación de la enzima <i>PA2125</i> con NAD(P) <sup>+</sup> .	86
Figura 18. Patrones de velocidad inicial de la reacción de la enzima <i>PA2125</i> con BZAL.	88
Figura 19. Cursos temporales de la actividad de la enzima <i>PA2125</i> en ensayos de cinética rápida.	90
Figura 20. Monómero del modelo por homología obtenido para la enzima <i>PA2125</i> .	94
Figura 21. Dímero y tetramero del modelo por homología obtenido para <i>PA2125</i> .	96
Figura 22. Sitio activo del modelo por homología obtenido para la enzima <i>PA2125</i> .	98
Figura 23. Sitio de unión de la coenzima del modelo por homología de la enzima <i>PA2125</i> .	101
Figura 24. Acoplamiento molecular (“docking”) de aldehídos aromáticos en el sitio activo de la enzima <i>PA2125</i> .	104
Figura 25. Acoplamiento molecular (“docking”) de TMABAL y 2-hidroxibenzaldehído en el sitio activo de la enzima <i>PA2125</i> .	106
Figura 26. Comparación de contextos genómicos de <i>PA2125</i> con otras ALDHs cuyo sustratos lo son también de <i>PA2125</i> .	112

## ÍNDICE DE ESQUEMAS

Esquema 1. Reacción general de las ALDHs hidrolíticas.	20
Esquema 2. Mecanismo químico de la reacción de las ALDHs hidrolíticas.	23
Esquema 3. Tipos de enzimas aldehído deshidrogenasas.	24
Esquema 4. Mecanismo cinético propuesto para las ALDHs hidrolíticas.	25
Esquema 5. Esquema general de las dos etapas del mecanismo químico de las ALDHs.	34
Esquema 6. Rutas metabólicas en las que participan los aldehídos que in vitro son sustratos de la PA2125.	110
Esquema 7. Estructuras resonantes de aldehídos aromáticos con grupos electrodonadores y electroatrayentes.	116

## ÍNDICE DE TABLAS

Tabla 1. Enzimas pertenecientes a la superfamilia ALDH en <i>P. aeruginosa</i> cepa PAO1.	61
Tabla 2. Porcentaje de identidad con <i>PA2125</i> de enzimas seleccionadas de su misma familia y otras enzimas ALDHs que poseen la actividad asignada a algunas de las enzimas de la familia de la <i>PA2125</i> .	64
Tabla 3. Resultados de la purificación de la enzima <i>PA2125</i> .	79
Tabla 4. Parámetros cinéticos aparentes de la enzima <i>PA2125</i> obtenidos con benzaldehído y aldehídos derivados.	85
Tabla 5. Parámetros cinéticos aparentes de la enzima <i>PA2125</i> obtenidos con NAD <sup>+</sup> y NADP <sup>+</sup> .	86
Tabla 6. Parámetros cinéticos de la enzima <i>PA2125</i> obtenidos con BZAL como sustrato.	89
Tabla 7. Parámetros cinéticos del estado pre-estacionario de la enzima <i>PA2125</i> .	91
Tabla 8. Constantes de sustitución para y meta para grupos sustituyentes en el anillo de benceno obtenidas mediante la ecuación de Hammett.	118

**ABREVIATURAS**

<b>ABAL</b>	4-Aminobutiraldehído
<b>ALDHs</b>	Aldehído deshidrogenasas
<b>AMADHs</b>	Aminoaldehído deshidrogenasas
<b>APAL</b>	3-Aminopropionaldehído
<b>BADH</b>	Betaína aldehído deshidrogenasa
<b>BAL</b>	Betaína aldehído
<b>BLAST</b>	Basic Local Alignment Search Tool (Herramienta Básica de Alineamiento Local)
<b>BZAL</b>	Benzaldehído
<b>CDD</b>	Conserved Domain Database (Base de datos de dominios conservados)
<b>CoA</b>	Coenzima A
<b>CTAB</b>	Bromuro de hexadecil trimetilamonio
<i>E. coli</i>	<i>Escherichia coli</i>
<b>EC</b>	Enzyme Commission
<b>GB</b>	Glicina betaína
<b>GMC</b>	Glucosa-metanol-colina
<i>HsALDH2</i>	Aldehído deshidrogenasa familia 2 de <i>Homo sapiens</i>
<b>IPTG</b>	Isopropil- $\beta$ -D-tiogalactopiranósido
<b>LB</b>	Medio Luria Bertani
<b>2-METOXIBZAL</b>	2-Metoxibenzaldehído
<b>NAD<sup>+</sup></b>	Dinucleótido de nicotinamida y adenina (forma oxidada)
<b>NADH</b>	Dinucleótido de nicotinamida y adenina (forma reducida)
<b>NADP<sup>+</sup></b>	Dinucleótido de nicotinamida y adenina fosfato (forma oxidada)
<b>NCBI</b>	Centro Nacional de Información Biotecnológica
<b>3-NITROBZAL</b>	3-Nitrobenzaldehído
<b>4-NITROBZAL</b>	4-Nitrobenzaldehído
<b>NMDPR</b>	The National Microbial Pathogen Database Resource (Recurso de Datos Nacional de Patógenos Microbianos)
<i>P. aeruginosa</i>	<i>Pseudomonas aeruginosa</i>



<b>PaBADH</b>	Betaína aldehído deshidrogenasa de <i>Pseudomonas aeruginosa</i>
<b>pET28b(+)-Pa2125</b>	Vector de expresión conteniendo el gen pa2125
<b>PCR</b>	Reacción en cadena de la polimerasa
<b>PDB</b>	Protein Data Bank
<b><i>P. putida</i></b>	<i>Pseudomonas putida</i>
<b>RAST</b>	Rapid Annotations using Subsystems Technology (Anotación rápida usando tecnología de subsistema)
<b>SAAL</b>	Salicilaldehído (2,5-dihidroxibenzaldehído)
<b>SaBADH</b>	Betaína aldehído deshidrogenasa de <i>Staphylococcus aureus</i>
<b>SoBADH</b>	Betaína aldehído deshidrogenasa de <i>Spinacia oleracea</i>
<b>SmGADP</b>	Gliceraldehído 3-fosfato deshidrogenasa no fosforilante de <i>Streptococcus mutans</i>
<b>SDS</b>	Dodecil sulfato de sodio
<b>SDS-PAGE</b>	Electroforesis desnaturalizante en gel de poliacrilamida en presencia de SDS
<b>SEC</b>	Size exclusion chromatography (Cromatografía de exclusión molecular)
<b>TMABAL</b>	4-Trimetilaminobutiraldehído

## RESUMEN

*Pseudomonas aeruginosa* es un patógeno humano oportunista, resistente a muchos antibióticos y desinfectantes. Esta bacteria posee la enzima betaína aldehído deshidrogenasa (*PA5373*, *PaBADH*) que participa en el segundo paso del catabolismo de colina consistente en la oxidación de betaína aldehído a glicina betaína, un eficiente osmoprotector que le permite a la bacteria crecer en las condiciones de estrés osmótico que imperan en los tejidos que infecta. Por esta razón se ha propuesto a la *PaBADH* como un posible blanco de fármacos contra este patógeno. La enzima *PA2125* fue anotada como *BADH* y si esta anotación es correcta también oxidaría a *BAL* y podría suplir a la *PaBADH* en caso de que esta última fuera inhibida. Sin embargo, la *PA2125* aún no ha sido caracterizada cinética o estructuralmente y por tanto se desconoce su especificidad por el sustrato aldehído. En este trabajo de tesis nos propusimos estudiar esta especificidad, lo que además nos permitiría inferir su papel fisiológico. Para ello, como primera aproximación, buscamos las proteínas pertenecientes a la misma familia que la *PA2125* dentro de la superfamilia de las aldehído deshidrogenasas (*ALDHs*). De las 179 encontradas, todas ellas bacterianas, sólo de dos se conoce su actividad: una oxida al furfural y la otra al 2,5-dihidroxibenzaldehído (gentisalaldehído), aunque encontramos otras varias también anotadas como *BADHs* pero no caracterizadas. Para estudiar experimentalmente la especificidad de la *PA2125*, clonamos y sobre-expresamos el gen *pa2125*, obtuvimos la enzima recombinante y la purificamos y caracterizamos cinéticamente. Encontramos que usa tanto a  $\text{NAD}^+$  como a  $\text{NADP}^+$  como cofactor, aunque prefiere al primero, que no usa a *BAL* o furfural como sustrato y que oxida sustratos aromáticos como benzaldehído o derivados. De estos, el 4-nitrobenzalaldehído es el mejor, seguido por el 3-nitrobenzalaldehído, 2-hidroxibenzaldehído (salicilaldehído), benzaldehído y 4-metoxibenzaldehído, en este orden.

Estos resultados sugieren la participación de la *PA2125* en alguna ruta de degradación de compuestos aromáticos contaminantes de suelos que tenga como intermediario a un aldehído aromático sustituido, pero no podemos decir por ahora con certeza cuál es este aldehído ni, por tanto, cuál es el compuesto o compuestos contaminantes en cuya degradación participaría.

En estudios de velocidad inicial con benzaldehído como sustrato encontramos que sigue un mecanismo cinético Bi Bi ordenado en estado estacionario, siendo el primer sustrato el cofactor y el segundo el aldehído. La afinidad aparente por ambos disminuye cuando aumenta la concentración del otro. En estudios de cinética rápida encontramos que el paso limitante de la reacción es la hidrólisis del intermediario tioéster o la liberación del cofactor reducido.

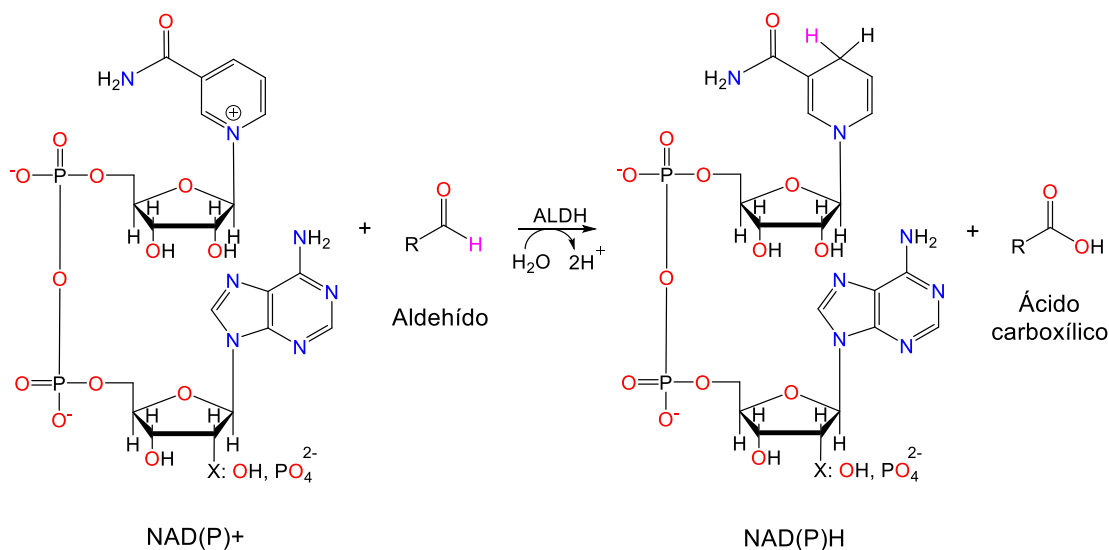
Se obtuvo un modelo por homología de la estructura tridimensional de la enzima con el que se realizaron acoplamientos moleculares con  $\text{NAD(P)}^+$  y con los aldehídos que fueron probados *in vitro*. El residuo Gly175 en el sitio de unión al cofactor le permite unir tanto a  $\text{NAD}^+$  como a  $\text{NADP}^+$ , pero no pudimos explicar la preferencia de la enzima *PA2125* por  $\text{NAD}^+$ . En cuanto a la especificidad por el aldehído, los resultados obtenidos indican que Ser150, Arg157, Phe281 y Glu449 son importantes para determinar esta especificidad. Esto se apoya además en los análisis de los contextos genómicos de los genes que codifican para las enzimas de la familia ALDH a la que pertenece *PA2125*, familia que parece estar formada al menos por cuatro subfamilias.

## I ANTECEDENTES

### I.1 Generalidades de las enzimas aldehído deshidrogenasas

Los aldehídos son compuestos orgánicos muy reactivos, en los cuales el carbono carbonílico es muy electrofílico debido a la deficiencia electrónica producida por la alta electronegatividad del oxígeno que forma el doble enlace con el carbono, de tal manera que el carbono carbonílico puede ser atacado por nucleófilos ( $\text{OH}^-$ ,  $\text{RO}^-$ ,  $\text{R-NH}_2$ ,  $\text{CN}^-$ ,  $\text{R-S}^-$ , etc.) para llevar a cabo reacciones de adición nucleofílica. Cuando el nucleófilo que ataca al carbono carbonilo es un grupo hidroxilo ( $-\text{OH}$ ) se produce un compuesto hemiacetal, pero cuando el nucleófilo es un grupo tiol ( $-\text{SH}$ ) se produce un tiohemiacetal. Biológicamente los aldehídos son compuestos intermediarios que se producen de forma exógena como endógena. Los aldehídos endógenos se producen en el metabolismo de lípidos, esteroides, carbohidratos, aminoácidos, aminos biogénicas y vitaminas (Vasiliou *et al.*, 2000). Al ser compuestos altamente reactivos por su carácter electrofílico, los aldehídos son tóxicos para las células incluso a bajas concentraciones, pero existen varios tipos de enzimas que los convierten en alcoholes o ácidos no tóxicos. Entre estas enzimas destacan las aldehído deshidrogenasas (ALDHs; EC 1.2.1) que catalizan la oxidación irreversible dependiente de nucleótido de nicotinamida y adenina ( $\text{NAD}^+$ ) o el nucleótido de nicotinamida y adenina fosfato ( $\text{NADP}^+$ ) (Jakoby, 1963) de una gran variedad de aldehídos alifáticos y aromáticos a sus respectivos ácidos carboxílicos (Lindahl, 1992; Yoshida *et al.*, 1998) (Esquema 1). Este tipo de enzimas están presentes en todos los organismos de la escala filogenética tanto en procariotas como eucariotas (Sophos y Vasiliou, 2003), y se ubican en todas las regiones subcelulares incluyendo el citosol (; Moore *et al.*, 2009; Singh *et al.*, 2013; Krupenko *et al.*, 2015), las mitocondrias (Liu y Schnable, 2002), peroxisomas, el retículo endoplasmático

(Ashibe *et al.* 2007) y el núcleo (Stagos *et al.*, 2010), con varias de estas enzimas encontradas en más de un compartimiento. En bacterias estas enzimas se encuentran solubles en el citosol.



**Esquema 1. Reacción general de las ALDHs hidrolíticas.** X es -OH en el caso del cofactor  $\text{NAD}^+$  o  $-(\text{PO}_4)^{2-}$  cuando el cofactor es  $\text{NADP}^+$ . R puede ser una cadena alifática o un anillo aromático.

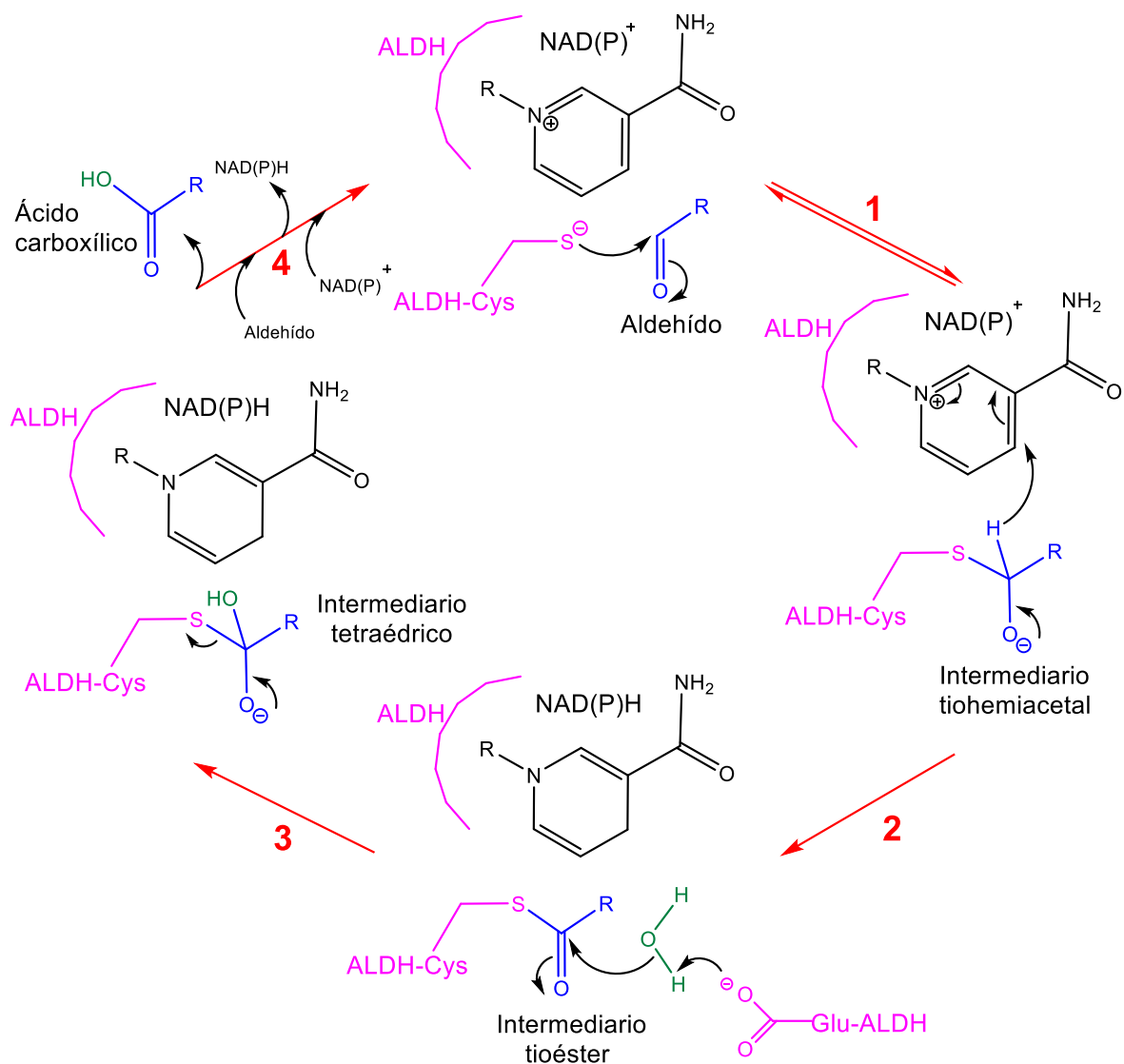
Las ALDHs se han clasificado por la Enzyme Commission (EC) con base en su especificidad por el sustrato aldehído, lo que sugiere su papel fisiológico. Sin embargo, en ocasiones esta clasificación no es precisa, porque a una misma enzima se le han asignado diferentes números EC dependiendo del sustrato con el que se haya medido su actividad. Tal es el caso de las betaína aldehído deshidrogenasas (BADHs) de animales que pueden usar como sustratos, además de la betaína aldehído (BAL), a diferentes  $\omega$ -aminoaldehídos y por ello aparte de su clasificación inicial como la isoenzima E3 de las aldehído deshidrogenasas (EC 1.2.1.5), se las ha clasificado también como la 4-N-trimetilaminobutiraldehído deshidrogenasa (EC 1.2.1.47) y la 4-aminobutiraldehído

deshidrogenasa (EC 1.2.1.19) (Muñoz-Clares *et al.*, 2010). El gran número de ALDHs cuya secuencia de aminoácidos ya se conoce, junto con las herramientas que hoy existen de comparación y análisis de secuencias, han permitido clasificarlas de acuerdo con sus relaciones filogenéticas en familias que agrupan proteínas que poseen al menos un 40% de identidad entre ellas y que en conjunto forman la superfamilia de las ALDHs (Sophos y Vasiliou, 2003; Black y Vasiliou, 2009; Riveros-Rosas *et al.*, 2013). Algunas ALDHs presentan una especificidad alta por los aldehídos que son sus sustratos, mientras que otras usan una amplia gama de aldehídos con estructuras moleculares parecidas (Riveros-Rosas *et al.*, 2013). El Centro Nacional de Información Biotecnológica (NCBI, por sus siglas en inglés) tiene la base de datos de dominios conservados (CDD, por su sigla en inglés) para la clasificación de secuencias de proteínas de acuerdo con huellas de dominios conservados e identifica de esta forma sitios funcionales en las proteínas; así mismo incluye estructuras de proteínas ya cristalizadas y las relaciona con su función (Marchler-Bauer *et al.*, 2011). En el caso de las ALDHs el CDD ha logrado relacionar los dominios conservados con las familias de ALDHs y de esta manera ha asignado una función a un dominio conservado de secuencia. El CDD asigna un código “cl” a una superfamilia de proteína y a una familia de proteínas con un código “cd”. A la superfamilia de ALDHs dependientes de NAD(P)<sup>+</sup> el CDD le ha asignado el código cl11961. Esta superfamilia agrupa a casi 60 familias, todas ellas con diferentes códigos cd, de las cuales 27 ya han sido numeradas y se conocen su especificidad y función fisiológica; el resto no han sido caracterizadas, por lo que aún son de función desconocida. Entre estas últimas se encuentra la familia a la que pertenece la PA2125, enzima ALDH que es objeto de estudio en esta tesis, como describiremos más adelante.

Algunas ALDHs presentan una especificidad alta por los aldehídos que son sus sustratos, mientras que otras usan una amplia gama de aldehídos con estructuras moleculares parecidas (Riveros-Rosas *et al.*, 2013). Teniendo en cuenta la diversidad de aldehídos endógenos o exógenos que pueden ser sus sustratos, las ALDHs cumplen muchas diferentes e importantes funciones fisiológicas. Existen enzimas ALDHs dependientes de  $\text{NAD}^+$ , otras dependientes de  $\text{NADP}^+$  y aquellas con especificidad dual, es decir que no presentan preferencia por uno u otro cofactor. Hasta la fecha la mayor cantidad de ALDHs que se han caracterizado son  $\text{NAD}^+$  específicas, en menor cantidad las  $\text{NADP}^+$  específicas, como la enzima gliceraldehído 3-fosfato deshidrogenasa no fosforilante de *Streptococcus mutans* (SmGADP, ALDH11) (Cobessi *et al.*, 1999), y en cantidad muy reducida las que usan de forma indistinta y eficientemente uno u otro cofactor, como es el caso de la betaína aldehído deshidrogenasa de *P. aeruginosa* (PaBADH, ALDH9) (Velasco-García *et al.*, 2000).

### **I.2 Mecanismo químico de reacción catalizada por las aldehído deshidrogenasas**

Esto se produce vía formación de un intermediario covalente (tiohemiacetal) entre el aldehído y la cisteína catalítica de la ALDH. El mecanismo químico propuesto para la reacción catalizada por las ALDHs consiste en una catálisis covalente y ácido-básica que consta de cuatro pasos principales (Feldman y Weiner, 1972) (Esquema 2): (1) Una vez que se han unido tanto el cofactor como el sustrato aldehído al sitio activo, el grupo tiol de la cisteína catalítica realiza un ataque nucleofílico sobre el carbono del grupo carbonilo del sustrato para formar un intermediario tiohemiacetal tetraédrico.

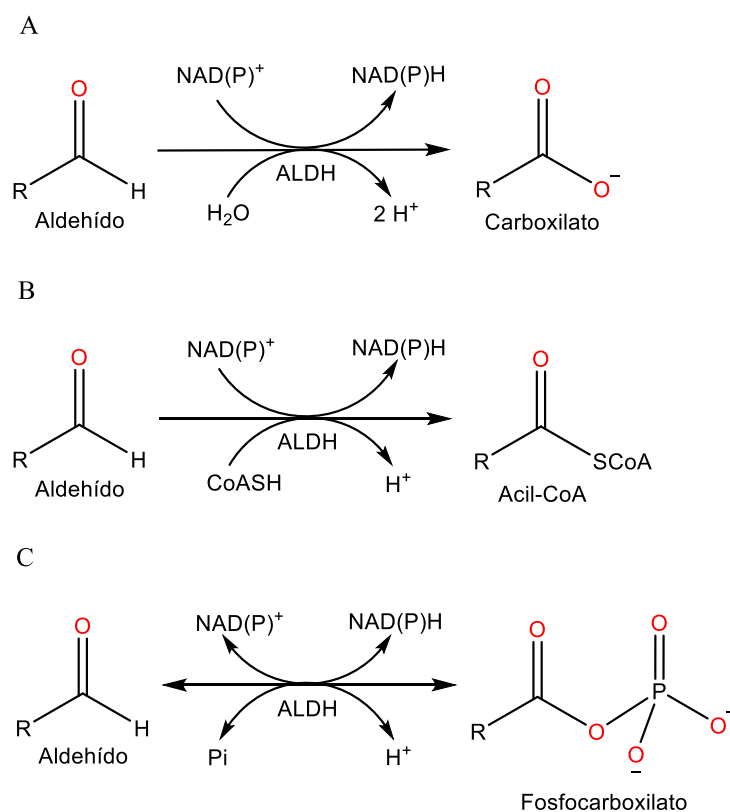


**Esquema 2. Mecanismo químico de la reacción de las ALDHs hidrolíticas.** Adaptado de Muñoz-Clares *et al.* (2011).

El oxianión formado en este paso es estabilizado mediante dos puentes de hidrógeno que recibe de la asparagina catalítica y del nitrógeno del grupo amida de la cadena principal de la cisteína catalítica. (2) Debido a la inestabilidad del intermediario tetraédrico, se regenera el enlace doble entre el oxígeno y el carbono carbonílico por el traslape de orbitales p de cada átomo, pasando de una geometría de hibridación  $sp^3$ , inestable en este caso, a una  $sp^2$  más estable, lo cual propicia la transferencia del hidruro desde el aldehído al cofactor



NAD(P)<sup>+</sup> para formar el intermediario tioéster y el cofactor reducido. (3) En el caso de las ALDHs hidrolíticas, la hidrólisis del tioéster por un ión hidroxilo, resultante de la activación de una molécula de agua (desprotonación) por el glutámico catalítico, libera el grupo tiol y el producto ácido de la reacción. (4) Por último, los productos salen del sitio activo de la enzima en forma ordenada, primero el ácido y después el cofactor reducido, regenerándose la enzima libre. Ni las enzimas ALDHs acilantes ni las fosforilantes poseen el residuo de glutámico que se requiere para desprotonar el agua catalítica para la hidrólisis del intermediario tetraédrico tioéster en la ALDHs hidrolíticas (Nagradova, 2001; D'Ambrosio *et al.*, 2006).

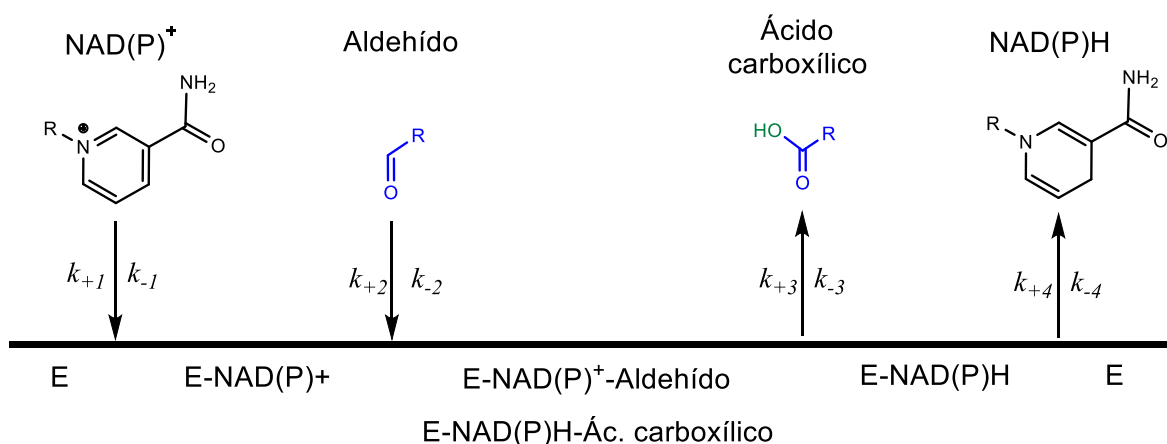


**Esquema 3. Tipos de enzimas aldehído deshidrogenasas.** (A) ALDHs hidrolíticas; (B) ALDHs acilantes; (C) ALDHs fosforilantes.

Atendiendo al criterio de la molécula nucleofílica que ataca al intermediario tioéster, existen tres tipos de ALDHs: hidrolíticas, en las que esta molécula es el agua, acilantes, en las que esta molécula es la coenzima A (CoA), y fosforilantes, en las que esta molécula es un fósforo inorgánico (Esquema 3).

### I.3 Mecanismo cinético

Las ALDHs hidrolíticas estudiadas hasta la fecha presentan un mecanismo cinético de tipo Bi Bi ordenado en estado estacionario (Esquema 4) (Feldman & Weiner, 1972). El primer sustrato en unirse a la enzima es el cofactor  $\text{NAD(P)}^+$  y el último producto en disociarse de la enzima es el  $\text{NAD(P)H}$ . Aunque el mecanismo cinético propuesto es consistente con el mecanismo químico para las reacciones catalizadas por este tipo de ALDHs, se desconoce la razón por la cual el nucleótido tiene que unirse antes que el aldehído, siendo que los sitios de entrada de ambos están en lados opuestos de cada subunidad, lo cual sugiere que podría llevarse una unión del sustrato y del cofactor al azar.



**Esquema 4. Mecanismo cinético propuesto para las ALDHs hidrolíticas.**

## **I.4 Propiedades estructurales de las aldehído deshidrogenasas**

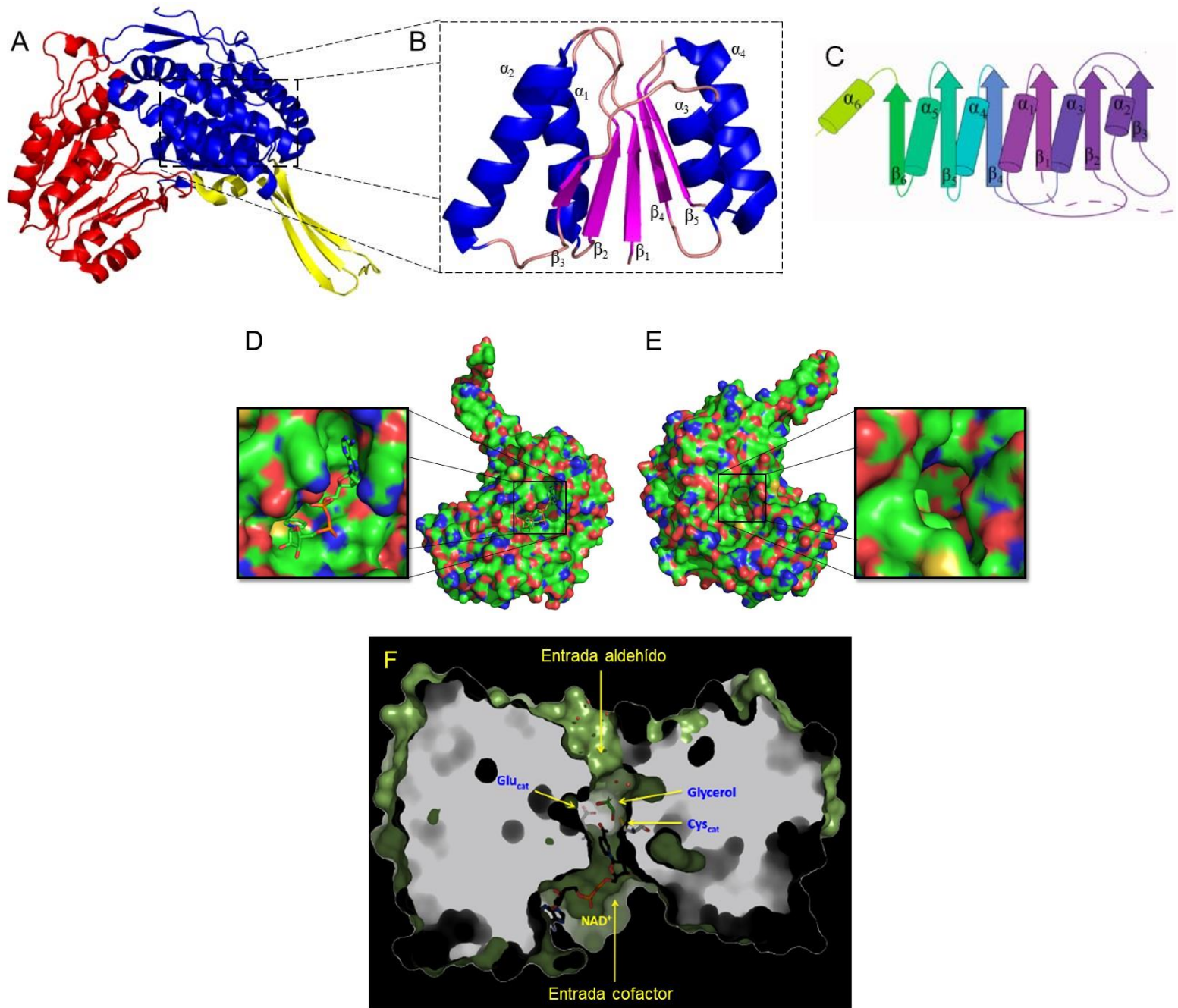
### ***I.4.1 Estructura primaria***

Las ALDHs poseen alrededor de 400 a 500 aminoácidos, de los cuales los residuos Cys y Asn catalíticos están absolutamente conservados (Hempel *et al.*, 1993). En las ALDHs hidrolíticas también está absolutamente conservado un residuo de glutámico, que juega igualmente un papel catalítico. Los principales residuos que participan en la determinación de la especificidad son el residuo en la posición 195 para el cofactor (González-Segura *et al.*, 2015) y los residuos en las posiciones 170, 177 y 465 para el aldehído (Riveros-Rosas *et al.*, 2013) (numeración de la ALDH2 humana, *HsALDH2*).

A octubre del 2017, el número de secuencias de genes de ALDHs disponibles en la base de datos GenBank (<http://www.ncbi.nlm.nih.gov/genbank/>) es de 632,073, de las cuales 166 son de virus, 587,798 de bacterias, 2,811 de arqueas, 2,126 de protistas, 13,309 de hongos, 9,005 de plantas y 16,858 de animales.

### ***I.4.2 Plegamiento de los monómeros***

Hasta esta fecha se han depositado en el Protein Data Bank (PDB; Burley *et al.*, 2017) 96 estructuras tridimensionales de diferentes ALDHs; todas ellas presentan un plegamiento similar para la unidad monomérica sin importar su estado de oligomerización. En cada una de las unidades monoméricas de las ALDHs se pueden diferenciar tres dominios: un dominio de unión al cofactor en la región N-terminal, un dominio catalítico y un dominio de oligomerización en la región C-terminal (Figura 1A).



**Figura 1. Estructura tridimensional de las enzimas aldehído deshidrogenasas.** (A) Representación en forma de lazos del monómero de la BADH de *P. aeruginosa* PaBADH (código PDB 2WME) mostrando el dominio catalítico (rojo), el dominio de unión al cofactor (azul) y el dominio de oligomerización (amarillo). (B) del Dominio de unión al cofactor en la BADH de *P. aeruginosa* PaBADH (código PDB 2WME) mostrando el plegamiento tipo Rossmann modificado, característico de las ALDHs, que carece del último motivo  $\alpha\beta$  del plegamiento Rossmann típico. Se muestran las

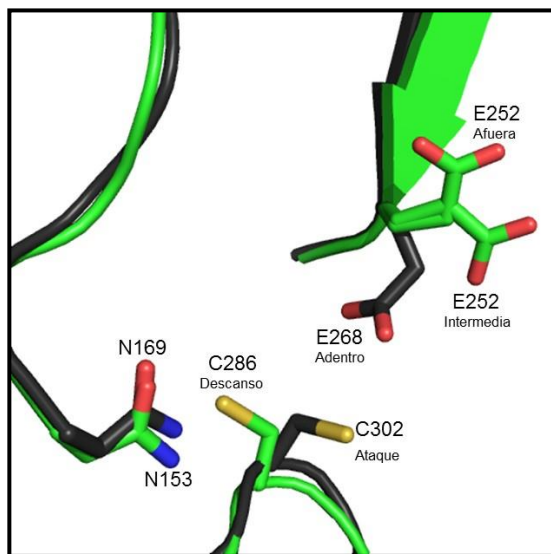
hebras  $\beta$  en color púrpura, las hélices  $\alpha$  en azul y las asas en rosa, (C) Plegamiento Rossmann encontrado en ALDHs en el que se muestran los motivos de hebra( $\beta$ )-asa-hélice( $\alpha$ ); (D) Sitio de unión al cofactor en PDB 2WME; (E) Túnel de entrada del aldehído en PDB 2WME. (F) Corte transversal de la estructura cristalográfica de la enzima betaína aldehído deshidrogenasa de *Spinacia oleracea* (PDB 4AOM) donde se muestra el túnel de entrada del aldehído y el túnel de entrada del cofactor NAD<sup>+</sup>.

*Dominio de unión al cofactor:* Posee un plegamiento tipo Rossmann modificado (Figura 1B), debido a que carece del último motivo  $\alpha\beta$  del plegamiento Rossmann típico (Rossmann *et al.*, 1974). Este tipo de plegamiento se encuentra en enzimas que tienen sitios de unión a cofactores como FAD, NAD<sup>+</sup> y NADP<sup>+</sup> y se caracteriza por estar compuesto por hebras  $\beta$  paralelas unidas a  $\alpha$  hélices siguiendo el orden topológico  $\beta\alpha\beta\beta$  (Hanukoglu, 2015). Como se especificó anteriormente las ALDHs usan al cofactor NAD<sup>+</sup> o al NAD(P)<sup>+</sup>, sin embargo la especificidad por la coenzima está determinada por la complementariedad electrostática y geométrica entre la enzima y el mononucleótido de adenina, con el que establece numerosos puentes de hidrógeno. La especificidad por el cofactor está determinada, principalmente, por la naturaleza del residuo en la posición 195 (numeración de la ALDH2 humana, *HsALDH2*) (González-Segura *et al.*, 2015). Las enzimas ALDH conocidas con glutamato en la posición 195 son bastante numerosas y prefieren NAD<sup>+</sup>; la mayor parte de ellas son incapaces de usar el NADP<sup>+</sup> porque la cadena lateral del glutámico produce un impedimento estérico y electrostático para la unión del grupo 2'-fosfato del NADP<sup>+</sup>, al que además repele electrostáticamente. Pero la flexibilidad conformacional del glutamato puede permitir la unión de NADP<sup>+</sup> en algunas enzimas si el grupo carboxilo se mueve lejos del grupo 2'-fosfato, lo cual es posible sólo si hay un pequeño residuo neutro en la posición 224 (numeración de la *HsALDH2*). En esta

posición pueden también existir aspartato, glutamina, histidina, lisina, arginina, prolina, serina, treonina, valina, leucina, isoleucina, alanina o glicina. Generalmente, las ALDHs que tienen un residuo diferente a glutamato en la posición 195, se unen con menos afinidad al  $\text{NAD}^+$  que las que poseen glutamato, debido a que no pueden hacer un puente de hidrógeno con el grupo  $2''\text{-OH}$  de la ribosa de la adenosina. Varias de ellas se sabe que se unen a  $\text{NADP}^+$ , incluso con mayor afinidad que al  $\text{NAD}^+$ , como es el caso de enzimas con Ser o Thr que suelen ser específicas para  $\text{NADP}^+$ . Un residuo de alanina o glicina en esta posición no interferiría con la unión del grupo  $2'$ -fosfato del  $\text{NADP}^+$ . Por lo tanto, es probable que enzimas con Ala/Gly195 se unan mejor a  $\text{NADP}^+$  que  $\text{NAD}^+$ , dadas las pocas interacciones que la proteína haría con la ribosa. Hasta la fecha, ninguna de estas enzimas se ha caracterizado cinética ni estructuralmente.

- *Dominio catalítico*: Cada unidad monomérica en las ALDHs contiene un dominio catalítico, en donde se encuentran tres residuos que están conservados en todas las ALDHs hidrolíticas (Muñoz-Clares *et al.*, 2017), ya que participan en la reacción de oxidación del sustrato aldehído: la cisteína catalítica, el glutámico catalítico y la asparagina catalítica. El sitio catalítico de las ALDHs se caracteriza por tener dos cavidades, una por donde entra el cofactor (Figura 1D) y otra que corresponde al túnel de entrada del aldehído (Figura 1E); en el lugar donde estas dos cavidades convergen se encuentra la cisteína catalítica. La cisteína catalítica (posición 302 de la *HsALDH2*) pertenece a un asa de este dominio, el asa catalítica, que contiene un giro beta tipo I y cuya estructura está conservada en todas las ALDHs con estructura tridimensional conocida (Muñoz-Clares *et al.*, 2017). El grupo tiol de la cadena lateral de este residuo es el encargado de hacer el ataque nucleofílico al carbono carbonílico del aldehído, mientras que el nitrógeno amídico de la cadena principal

forma un puente de hidrógeno con el oxígeno del grupo carbonilo del aldehído, o con el oxianión del tiohemiacetal que se forma durante la reacción. De las tres conformaciones que un residuo de cisteína puede adoptar, la cisteína catalítica de las ALDHs sólo puede adoptar dos (Figura 2) (Muñoz-Clares *et al.*, 2017).



**Figura 2. Sitio activo de las aldehído deshidrogenasas mostrando los residuos catalíticos y las conformaciones posibles de la cisteína y el glutámico catalíticos.** Se muestran la conformación “de descanso” de la cisteína catalítica y las conformaciones “intermedia” y “afuera” del Glu252 catalítico observadas en la estructura cristalográfica de la BADH de *P. aeruginosa* (*PaBADH*, código PDB 2WME, cadena color verde). En posición de “ataque” se muestra a la Cys302 catalítica y al Glu268 catalítico en posición de “adentro” como se observó *HsALDH2* (código PDB 1O02, cadena color negra). La asparagina catalítica (Asn153 y Asn169 en la *PaBADH* y *HsALDH2*, respectivamente) siempre se ha encontrado en una sola conformación, que es la mostrada en esta figura para las dos enzimas.

En la primera, llamada conformación “de descanso” (González-Segura, *et al.*, 2009), la cadena lateral de la cisteína se encuentra dirigida hacia la cavidad del

cofactor, por lo cual el grupo tiol no puede realizar el ataque nucleofílico sobre el grupo carbonilo del aldehído. En la segunda posición, llamada conformación “de ataque” (González-Segura, *et al.*, 2009), la cadena lateral de la cisteína se encuentra dirigida hacia el sitio de unión del aldehído de tal manera que el grupo tiol pueda llevar a cabo el ataque nucleofílico sobre el aldehído y formar el intermediario tiohemiacetal en el paso 1 del mecanismo de reacción de las ALDHs

El glutámico catalítico (Glu268 de la HsALDH2) presente sólo en las ALDHs hidrolíticas es el residuo encargado de activar la molécula de agua para realizar el ataque nucleofílico al tioéster formado por la oxidación del tiohemiacetal. De acuerdo con lo observado en estructuras de ALDHs ya cristalizadas, la cadena lateral del residuo glutámico catalítico se puede encontrar en tres posibles conformaciones (González-Segura, *et al.*, 2009). La primera se la denominó “adentro” (Figura 2) debido a que el grupo carboxilo de la cadena lateral del residuo glutámico está cerca de la cisteína catalítica para arrancarle un protón y formar el tiolato, el cual es un mejor nucleófilo que el grupo tiol. La segunda conformación se denomina “intermedia” (Figura 2); en esta posición la cadena lateral del glutámico toma el protón de la molécula del agua hidrolítica, para de esta manera activarla para el ataque nucleofílico sobre el tioéster. A la tercera conformación se la llamó “afuera”, podría participar en la descarga del protón que la cadena lateral del glutámico tomó del agua hidrolítica con la ayuda de la Lys178 y Glu476 (numeración de la HsALDH2). La asparagina catalítica (Asn169, numeración de la ALDH2 humana) es el residuo cuyo grupo amida de su cadena lateral forma la cavidad del oxianión junto con el nitrógeno amídico de la cisteína catalítica,

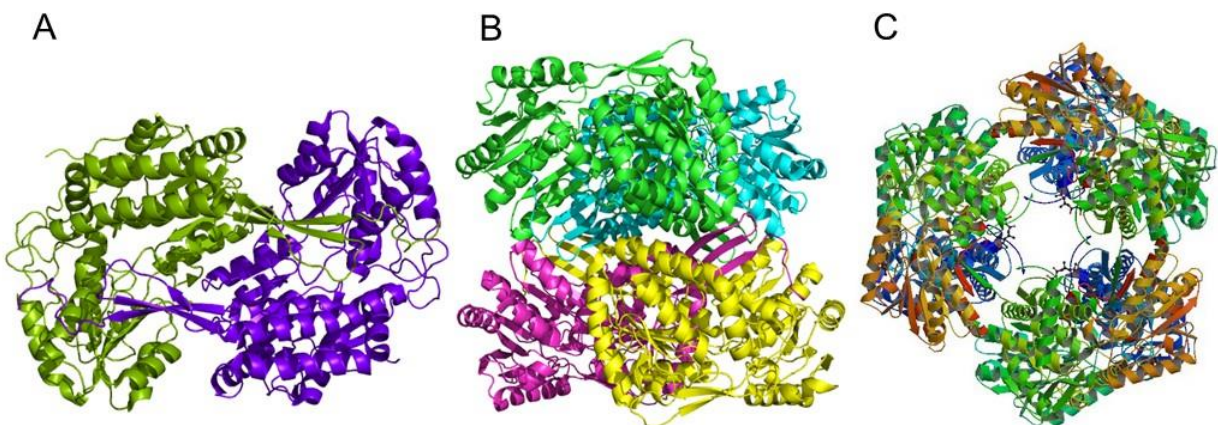


estabilizando por medio de puentes de hidrógeno al oxianión que se genera en el intermediario tiohemiacetal tetraédrico.

- *Dominio de oligomerización:* Este dominio es una extensión del dominio de unión al cofactor. Está constituido por tres o cuatro hebras  $\beta$  formando una hoja  $\beta$  que se extiende para entrar en contacto con una hoja  $\beta$  del dominio catalítico de otro monómero, lo cual permite la formación y estabilización de una unidad dimérica (Liu *et al.*, 1997; Johansson *et al.*, 1998).

### 1.4.3 Estructuras cuaternarias

Las ALDHs de estructura tridimensional conocida presentan estructuras cuaternarias diméricas, como es el caso de la estructura de la BADH de espinaca (*SoBADH*, código PDB: 4A0M) (Figura 3A), tetrámicas (dímero de dímeros), como en la BADH de *P. aeruginosa* (*PaBADH*, código PDB: 2WME) (Figura 3B) o hexaméricas (trímero de dímeros), en unos pocos casos como en la estructura de  $\Delta$ -1-pirrolina-5-carboxilato deshidrogenasa de *Thermus thermophilus* (código PDB: 1UZB) (Figura 3C).



**Figura 3. Tipos de estructuras cuaternarias encontradas en aldehído deshidrogenasas.**

(A) Dímero de la enzima BADH de espinaca (*SoBADH*, código PDB 4A0M). (B) Tetrámero (dímero de dímeros) de la BADH de *P. aeruginosa* (*PaBADH*, código PDB

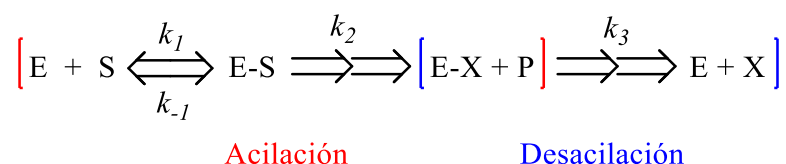
2WME). (C) Hexámero (trímero de dímeros) de la estructura de  $\Delta$ -1-pirrolina-5-carboxilato deshidrogenasa (glutamato semialdehído deshidrogenasa) de *Thermus thermophilus* (código PDB 1UZB). Figuras hechas en Pymol (<https://pymol.org/2/>).

### **I.5 Cinética del pre-estado estacionario de las aldehído deshidrogenasas**

En un experimento de cinética enzimática típica, la reacción inicia cuando se combinan la enzima y los sustratos libres. El período inmediatamente posterior a la iniciación de la reacción se caracteriza por el aumento en la concentración de los intermediarios que se producen a lo largo de la reacción y se denomina periodo pre-estacionario. Es en el estado pre-estacionario donde podemos observar y analizar el primer ciclo catalítico. El período de estado pre-estacionario es seguido por subsecuentes ciclos catalíticos durante los cuales estos intermedios de la reacción permanecen a una concentración relativamente constante; a este periodo se le conoce como estado estacionario.

El ciclo catalítico en las ALDHs está conformado por dos etapas principales: la etapa donde ocurre la acilación (formación del intermediario tioéster) y la etapa donde se lleva a cabo la desacilación (formación del producto ácido) (Esquema 5). La etapa de acilación comprende la unión no covalente entre la enzima (E), el cofactor  $\text{NAD(P)}^+$  (N) y el sustrato (S) para la formación del complejo ternario E.N.S, seguida de la formación del intermediario tiohemiacetal por el ataque nucleofílico de las cisteína catalítica al carbono carbonílico del aldehído, y la formación tanto del intermediario tioéster como del cofactor reducido  $\text{NAD(P)H}$ , que ocurre por la transferencia de hidruro del intermediario tiohemiacetal al anillo de nicotinamida del cofactor. No olvidar que para la etapa de desacilación ya se encuentra formado el cofactor reducido  $\text{NADP(H)}$ . La etapa de desacilación en las ALDHs

hidrolíticas comprende la formación de un intermediario tetraédrico (resultado del ataque nucleofílico por parte de una molécula de agua al intermediario tioéster), seguido de la ruptura del enlace C-S que da lugar a la formación del producto ácido; posteriormente se da la salida de este producto, seguido de la salida del cofactor reducido NAD(P)H.



**Esquema 5. Esquema general de las dos etapas del mecanismo químico de las ALDHs.**

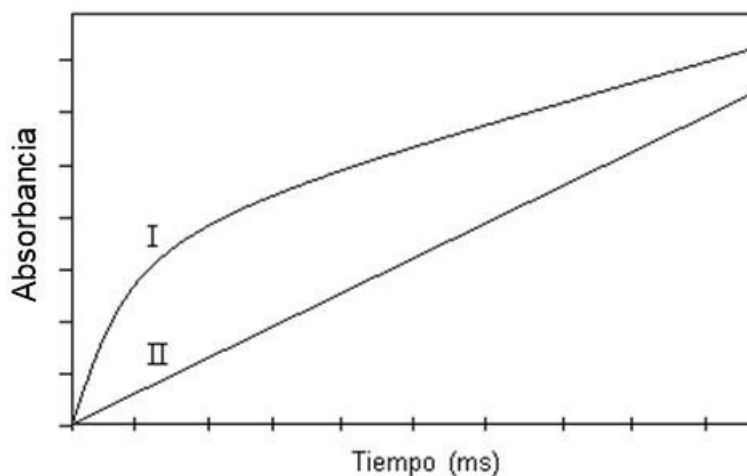
Entre corchetes rojos se muestra la etapa de acilación y entre corchetes azules la etapa de desacilación. E = complejo E-NAD(P)<sup>+</sup>; S = aldehído; E-S = complejo enzima-sustrato, no unido covalentemente; E-X= intermediario tioéster; P = NAD(P)H; X = producto ácido.

En el Esquema 5 se puede observar un esquema general de los pasos del mecanismo químico de las ALDHs con sus correspondientes constantes de velocidad:  $k'_1$  que es la constante de velocidad de pseudoprimer orden de acilación ( $k_{ac}$  en  $s^{-1}$ ) la cual está conformada por las constantes de velocidad  $k_1$ ,  $k_{-1}$  y  $k_2$  de acuerdo con la Ecuación 1, en donde  $K_d$  es la constante de disociación del sustrato del complejo E-S ( $K_d = k_{-1} / k_1$ ):

$$k'_1 = k_2[S]/([S] + K_d) \qquad \qquad \qquad \text{Ecuación 1}$$

$k_3$  es la constante de velocidad de primer orden de desacilación ( $k_{dac}$  en  $s^{-1}$ ). Si  $k'_1$  es mayor que  $k_3$ , se observará un “burst” o estallido en la producción de NAD(P)H, es decir, se observará una primera fase exponencial que corresponde a la formación de P durante el primer ciclo catalítico, seguida de una fase lineal, que corresponde a la velocidad de producción de P en los siguientes ciclos catalíticos, y que por tanto será igual a la velocidad

de la reacción durante el estado estacionario (Figura 4). Si  $k'_1$  es menor que  $k_3$  la velocidad observada desde el primer ciclo catalítico será la velocidad de estado estacionario



**Figura 4. Ejemplo de cursos temporales de enzimas aldehído deshidrogenasas en ensayos de cinética rápida.** Progresión de una reacción con “burst” en la formación de NAD(P)H (I) y sin “burst” en la formación de NAD(P)H (II).

Asumiendo condiciones de seudoprimer orden (es decir, que la  $[S]$  sea mucho mayor que la concentración de sitios activos  $[E_0]$ ), la producción de P a lo largo del tiempo sigue la siguiente ecuación:

$$[P] = [E_0] \left( \frac{k'_1}{k'_1 + k_3} \right) \times \left\{ \left( \frac{k'_1}{k'_1 + k_3} \right) (1 - e^{-(k'_1 + k_3)t}) + k_3 t \right\} \quad (\text{Ecuación 2})$$

La amplitud de la fase exponencial ( $A_{\text{burst}}$ ) está dada por:

$$A_{\text{burst}} = [E_0] \left( \frac{k'_1}{k'_1 + k_3} \right)^2 \quad (\text{Ecuación 3})$$

Donde  $A_{\text{burst}}$  será igual a  $[E_0]$  si  $k'_1$  es mucho mayor que  $k_3$ , y si este es el caso el valor de  $A_{\text{burst}}$  es igual a la concentración de sitios activos. Si esta condición no se cumple,  $A_{\text{burst}}$  tendrá un valor inferior a  $[E_0]$  y no se puede conocer la concentración de sitios activos por medio de un experimento de este tipo. Téngase en cuenta que  $A_{\text{burst}} = [E_0]$  siempre que  $k'_1$

$\gg k_3$ , aunque [S] no sea saturante. La constante de velocidad del “burst” ( $k_{burst}$ ) está dada por:

$$k_{burst} = k'_1 + k_3 \quad (\text{Ecuación 4})$$

Si  $k'_1 \gg k_3$ , la  $k_{burst}$  será igual a  $k'_1$  o  $k_{ac}$ ; de esta manera podemos conocer  $k'_1$ , que depende de la [S] tal como se dijo en la ecuación Ecuación 1.

La velocidad del estado estacionario ( $v_{ss}$ ) es la pendiente de la fase lineal y está dada por:

$$m = [E_0] k'_1 k_3 / (k'_1 + k_3) \quad (\text{Ecuación 5})$$

Luego si  $k'_1 \gg k_3$ ,  $m = [E_0] k_3$  y nos permite conocer  $k_3$ . Si esta condición no ocurre,  $m$  puede ser menor, igual o mayor a  $[E_0] k_3$ .

## **I.6 Género *Pseudomonas***

El género *Pseudomonas* perteneciente a la clase  $\gamma$ -proteobacterias y familia *Pseudomonadaceae* (Woese, 1987) está formado por un grupo diverso de bacilos Gram-negativos no esporulantes, que miden entre 0.5 – 1.0  $\mu\text{m}$  de diámetro por 1.5 – 5.0  $\mu\text{m}$  de longitud, presentan movilidad por uno o varios flagelos polares y son aeróbicos, aunque bajo condiciones anaeróbicas pueden usar nitrato como aceptor alternativo de electrones. Los miembros de este género son capaces de vivir y crecer en diversas condiciones ambientales, habitan en suelos (Cho y Tiedje, 2000) y ecosistemas acuáticos (Elomari *et al.*, 1996) y varios son patógenos de animales, incluidos humanos (Oberhardt *et al.*, 2008). y/o plantas (Rahme *et al.*, 1995; Raaijmakers *et al.*, 2002; Haas y Keel, 2003; Morris *et al.*, 2010). Algunas especies del género *Pseudomonas* son promotoras del crecimiento vegetal a través del control de microorganismos patógenos, produciendo compuestos antifúngicos o antibióticos que evitan enfermedades en las raíces de las plantas (Dowling y O’Gara, 1994).

También pueden actuar como solubilizadoras de fósforo por su capacidad de producir ácidos orgánicos (ácido oxálico, fumárico y cítrico) y enzimas fosfatasas, facilitando la solubilización del fósforo inorgánico (Ferraris y Couretot, 2004).

### ***1.6.1 Pseudomonas aeruginosa***

*Pseudomonas aeruginosa* es la especie más conocida e importante del género por su interés clínico, como un importante patógeno humano. Rara vez causa enfermedad en personas sanas, pero es un patógeno oportunista causando neumonía en pacientes en nosocomios (Bergmans *et al.*, 1998; Parker *et al.*, 2008; Fujitani *et al.*, 2011), serias infecciones en pacientes con quemaduras graves (Estahbanati *et al.*, 2002; Armour *et al.*, 2007; Ressler *et al.*, 2008), cateterizados del tracto respiratorio y urinario (Morris *et al.*, 1999; Mittal *et al.*, 2009), inmunocomprometidos como los enfermos de SIDA (Shepp *et al.*, 1994) o bajo quimioterapia (Rolston *et al.*, 1992; Bodey *et al.*, 2001). Es la principal causa de muerte en enfermos con fibrosis quística (Smith *et al.*, 1996; Nixon *et al.*, 2001; Worlitzch *et al.*, 2002). Presenta una alta resistencia a antibióticos como penicilinas (Piperacilina/tazobactam), cefalosporinas (Ceftazidime), quinolonas y quinolinas (ciprofloxacina), carbapenem (Imipenem, Meropenem), aminoglucósidos (Amikacina, Gentamicina, Tobramicina), glicopéptidos (Vancomicina), cloramfenicol, imidazólicos (Metronidazol) y sulfamidas (Lambert, 2002; Aloush *et al.*, 2006; Bredenstein *et al.*, 2011). *Pseudomonas aeruginosa* es uno de los patógenos oportunistas más temidos debido a su múltiple resistencia a fármacos, por lo cual ha recibido el denominativo "superbacteria" (Murray *et al.*, 2015) presentando limitadas opciones terapéuticas. La habilidad de *P. aeruginosa* para causar una amplia gama de infecciones en los seres humanos se debe a su capacidad para producir un gran repertorio de factores de virulencia y, en última instancia,

responder y adaptarse a condiciones adversas como las impuestas por la respuesta inmune del huésped y la exposición a los antibióticos (Pendleton *et al.*, 2013). De ahí la importancia de buscar nuevas dianas para el desarrollo de nuevos fármacos que conduzcan a tratamientos más eficaces en el combate a la infección por esta “superbacteria”.

### **I.7 Enzimas aldehído deshidrogenasas de *Pseudomonas aeruginosa* PAO1**

Según Sophos y Vasiliou (2002) en el genoma de *P. aeruginosa* cepa PAO1, que es la más usada en los estudios sobre esta bacteria, hay 26 genes que codifican enzimas ALDHs, de las cuales hasta la fecha se ha estudiado la *PaBADH* (PA5373), tanto su actividad enzimática y expresión (Nagasawa *et al.*, 1976; Velasco-García *et al.*, 1999; Velasco-García *et al.*, 2000; Velasco-García, *et al.*, 2006) como su estructura (González-Segura *et al.*, 2009; Díaz-Sánchez *et al.*, 2011; Garza-Ramos *et al.*, 2013) e inhibición por fármacos (Velasco-García *et al.*, 2003; Zaldívar-Machorro *et al.*, 2011). Hasta la fecha sólo se conoce la estructura tridimensional de dos de ellas; de la *PaBADH* (código PDB 2WME), y de la N-succinilglutamato 5-semialdehído deshidrogenasa (PA0898) (PDB 3JU8). Existen ocho ALDHs de *P. aeruginosa* a las que no se les ha asignado aún una función fisiológica, por lo que en el genoma están anotadas ya sea como aldehído deshidrogenasas o como probables aldehído deshidrogenasas. Al resto de enzimas se les asignó una actividad por comparación de secuencias de aminoácidos con ALDHs de otros organismos cuya función fisiológica es conocida.

## II JUSTIFICACIÓN DEL PROYECTO

La enzima PA2125 (acceso NCBI: NP\_250815) es una ALDH compuesta por 482 aminoácidos que posee una masa molecular teórica de 51.3 kDa y un punto isoelectrico teórico de 5.55 (ambos calculados en el portal <http://web.expasy.org/protparam>). Inicialmente esta enzima fue anotada en las bases de datos NCBI y otras bases de datos como betaína aldehído deshidrogenasa. Sin embargo, como se indicó anteriormente, *P. aeruginosa* posee una enzima betaína aldehído deshidrogenasa (*PaBADH*) bien caracterizada que participa en la degradación de colina y en al síntesis del osmoprotector glicina betaína. El gen que codifica para la *PaBADH* (*betB*) forma parte del operón *betIBA* en el que además están incluidos los genes que codifican para la colina deshidrogenasa (*betA*) y para el represor de la transcripción *BetI* (*betI*), que responde a colina y se une a una región específica del promotor de este operón (Smith *et al.*, 1988) y de operones similares en otras bacterias (Lamark *et al.*, 1996). Cuando hay colina en el medio, ésta se une al represor *BetI* permitiendo que se transcriba el operón. La expresión de la *PaBADH* depende por tanto casi en forma absoluta de la presencia de colina en el medio de crecimiento de la bacteria (Nagasawa *et al.*, 1976; Velasco-García *et al.*, 1999; Velasco-García *et al.*, 2000; Velasco-García *et al.*, 2003; Velasco-García *et al.*, 2006; González-Segura *et al.*, 2009), condición que se da en los sitios de infección. Sin embargo, no podemos descartar la posibilidad de exista otra ruta de degradación de colina, al menos hasta la formación de GB, adicional a la ya estudiada y en la que podría estar participando la enzima PA2125. Teniendo en cuenta que *P. aeruginosa* es un patógeno humano oportunista de interés clínico considerado como una “superbacteria”, para nuestro grupo de investigación el conocer cuáles enzimas están involucradas en esta otra ruta, en caso de existir, es fundamental para nuestros esfuerzos de encontrar un compuesto antibacteriano



que produzca por inhibición enzimática la acumulación de BAL, un aldehído tóxico (Rathinasabapathi *et al.*, 1996). Por esta razón, estudiamos la PA2125, para conocer si realmente es una BADH, es decir, si puede oxidar a BAL y por tanto pudiera complementar la actividad de la PaBADH en caso de que inhibiéramos a ésta. En caso de no ser una BADH, aún es interesante su estudio puesto que es una enzima ALDH que no ha sido caracterizada ni bioquímica ni cinéticamente a la fecha (ni ella, ni ninguna de las enzimas que pertenecen a la misma familia de las ALDHs) y un estudio preliminar sobre esta enzima realizado en nuestro grupo de trabajo sugirió la existencia en su sitio activo de residuos posibles determinantes de la especificidad por el aldehído muy poco frecuentes en las ALDHs en general. Por ello, su estudio aportará información valiosa para ampliar nuestro conocimiento de las ALDHs de *P. aeruginosa*, muy poco estudiadas a la fecha y algunas de las cuales pudieran ser blanco de fármacos o de interés en biorremediación, y en general de los mecanismos moleculares por los cuales se determina la especificidad por el sustrato aldehído en la superfamilia de las ALDHs.

### III HIPÓTESIS

La enzima PA2125 de *P. aeruginosa* no oxida a betaína aldehído, a pesar de que fue anotada como betaína aldehído deshidrogenasa, y por tanto: 1) no puede suplir a la PaBADH en caso de que ésta esté inhibida, y 2) y su función fisiológica, aún desconocida, es diferente a la de la degradación de colina o síntesis del osmoprotector glicina betaína.

## IV OBJETIVOS

### IV.1 Objetivo general

Estudiar la especificidad por el aldehído de la enzima *PA2125* de *P. aeruginosa* e inferir su posible función fisiológica.

### IV.2 Objetivos particulares

- 1) Investigar la especificidad de *PA2125* por el sustrato aldehído mediante alineamientos múltiples de su secuencia de aminoácidos con las de posibles proteínas ortólogas.
- 2) Clonar el gen *pa2125* para obtener la enzima *PA2125* de forma recombinante y purificarla a homogeneidad.
- 3) Estudiar la especificidad de la enzima *PA2125* por el sustrato aldehído a concentraciones fisiológicas del cofactor determinando los parámetros cinéticos aparentes ( $V$ ,  $K_m$  y  $V/K_m$ ) para aquellos aldehídos que puedan ser oxidados por esta enzima.
- 4) Determinar el mecanismo cinético de la enzima con el aldehído sustrato que se considere de interés.
- 5) Conocer la estructura tridimensional de la enzima *PA2125* construyendo un modelo por homología.
- 6) Explorar las bases estructurales de la especificidad de la enzima *PA2125* por los aldehídos probados a través de estudios de acoplamiento molecular (“*docking*”) en el sitio activo del modelo por homología de la enzima *PA2125*.

## V MATERIALES Y MÉTODOS

### V.1 Estrategia experimental

La estrategia experimental seguida durante el desarrollo de este proyecto de tesis se esquematiza en la Figura 4.

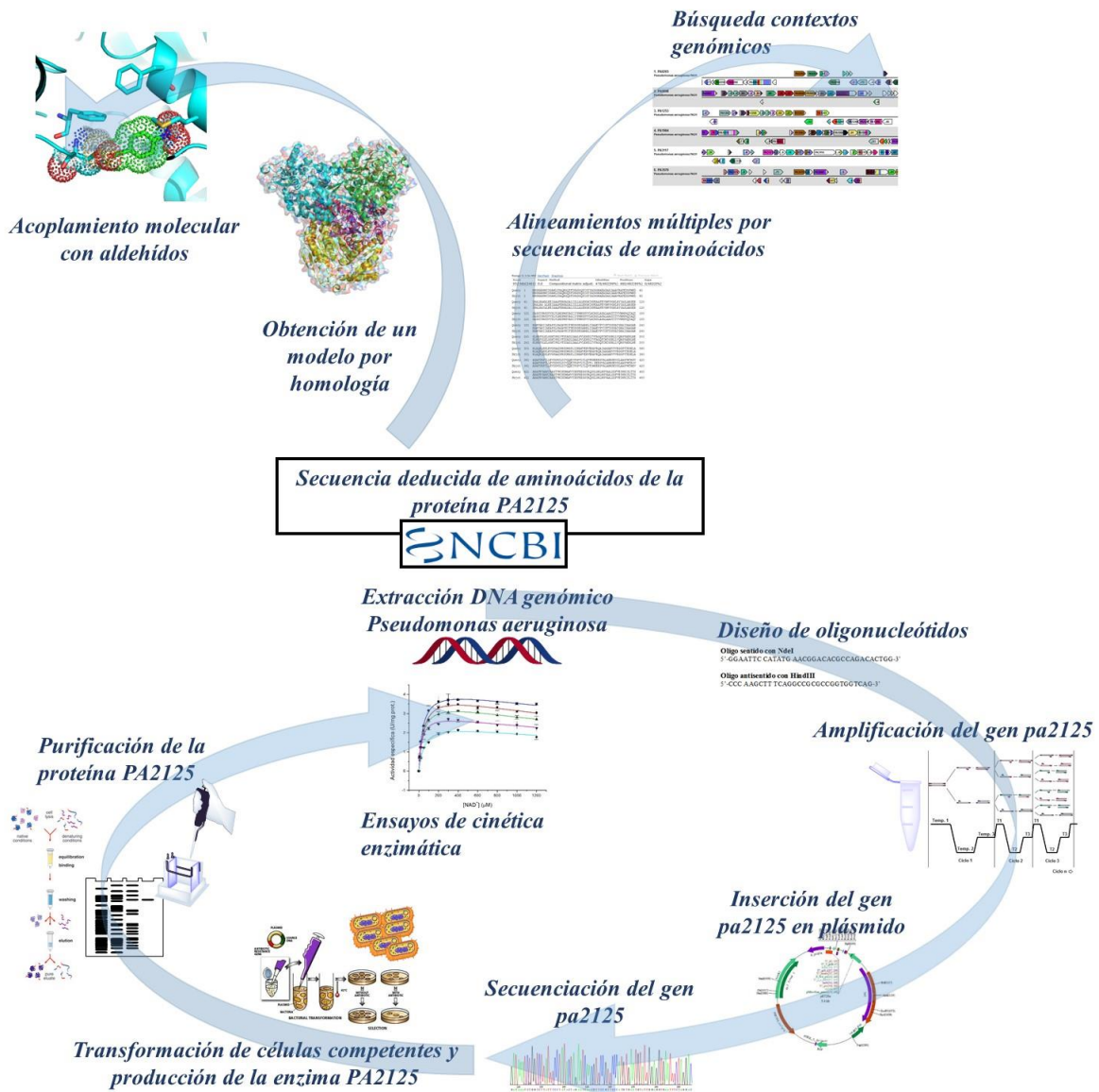


Figura 5. Estrategia experimental planteada para el desarrollo de la investigación de este trabajo de tesis.

## V.2 Búsqueda de ortólogos y alineamientos múltiples de secuencias de aminoácidos

La secuencia de aminoácidos de la enzima PA2125 del genoma de *P. aeruginosa* cepa PAO1 se obtuvo de las bases de datos de proteínas anotadas en NCBI con el código de acceso NP\_250815.1. Usando esta secuencia en formato FASTA como cebo se obtuvieron secuencias ortólogas de enzimas ALDH anotadas en el genoma de *P. aeruginosa* PAO1 en la base de datos del portal *Pseudomonas* Genome DB (<http://www.pseudomonas.com>; Winsor *et al.*, 2016)), empleando la herramienta en línea *Basic Local Alignment Search Tool* (BLAST). La búsqueda de proteínas ortólogas en organismos diferentes a *P. aeruginosa* se hicieron usando la herramienta BLAST del servidor de NCBI. A las proteínas que no pertenecen a una familia ALDH previamente numerada se les asignó una familia de acuerdo con la base de datos *Conserved Domain Database* (CDD; <http://www.ncbi.nlm.nih.gov/cdd>), porque esta clasificación de proteínas está de acuerdo con las familias de ALDH reportadas y con análisis filogenéticos realizados anteriormente realizados en nuestro grupo de investigación (Muñoz-Clares *et al.*, 2017). En el CDD a las superfamilias se les asigna códigos que comienzan con “cl”, y las familias se les asignan códigos que comienzan por “cd”. Para tratar de identificar la función de la proteína PA2125 de *P. aeruginosa* PAO1 se utilizó además un enfoque basado en el análisis del contexto genómico. En primer lugar, se descargó del NCBI el genoma completo de *P. aeruginosa* PAO1 (acceso NCBI: NC\_250815.1) y los genomas de los organismos que tienen enzimas ortólogas a PA2125 que se describen en la sección de Resultados. Después, el genoma fue re-anotado utilizando el servidor Anotación rápida usando tecnología de subsistema (RAST, por sus siglas en inglés) (Aziz *et al.*, 2008) que pertenece a la base de datos Recurso de Datos Nacional de Patógenos Microbianos (NMDPR por su siglas en inglés;

<http://www.nmpdr.org/FIG/wiki/view.cgi/Main/WebHome>). A continuación, utilizando la secuencia proteica de *PA2125* se realizó una búsqueda por BLAST para conocer los contextos genómicos de enzimas ortólogas a *PA2125* y compararlos con el contexto genómico de esta enzima. Los alineamientos múltiples de las secuencias de aminoácidos de proteínas seleccionadas, que se hicieron con diferentes fines y que describen en la sección de Resultados, se realizaron con la ayuda del programa computacional BioEdit (Hall, 1999), para la edición y manipulación de secuencias de aminoácidos, y ClustalX2 (Larkin *et al.*, 2007), para los alineamientos de las secuencias.

### **V.3 Clonación del gen *pa2125***

La secuencia de nucleótidos del gen *pa2125* se obtuvo del genoma de *P. aeruginosa* PAO1 reportado en la base de datos *National Center for Biotechnology Information* (NCBI), en la que este gen está consignado con el identificador *Gene ID*: 878590. Para la obtención del DNA genómico se cultivaron células de la cepa PAO1 en 10 mL de medio líquido LB durante 16 h a 37 °C y 200 rpm. Luego de este tiempo se tomaron alícuotas de 1,5 mL y se centrifugaron durante 4 min a 9,000 rpm (8,784 g); se desechó el sobrenadante y se adicionaron 567 µL de amortiguador TE (10 mM Tris, 1 mM EDTA, pH 8.0), 30 µL de dodecil sulfato sódico (SDS) al 10% y 3 µL de proteinasa K (Sigma-Aldrich, EUA) a una concentración de 20 mg/mL. La mezcla se incubó por 1 h a 37 °C. Posteriormente, se agregaron 100 µL de NaCl 5 M y 80 µL de una solución de bromuro de hexadeciltrimetilamonio (CTAB) 10% (p:v) / NaCl 0.7 M; se mezcló mediante agitación y se incubó por otros 10 min a 65 °C. A continuación, se agregó un volumen igual de cloroformo/alcohol isoamílico (24:1), se mezcló por agitación y se promovió la

precipitación del DNA dejando el tubo toda la noche a -20°C, tras lo cual se centrifugó durante 5 min a 9,000 rpm (8,784 g). Se desechó el sobrenadante y se realizaron varios lavados con etanol al 70% (v/v) centrifugando en cada uno de ellos a 9,000 rpm (8784 g) por 5 min y desechando el sobrenadante. Finalmente, la pastilla de DNA se resuspendió con 50 µL de agua estéril desionizada, la concentración de DNA se determinó mediante espectrofotometría, midiendo la absorbancia a 260 nm en un equipo *NanoDrop* 2000 serie F515 (Thermo Scientific, EUA), y se almacenó a -20 °C.

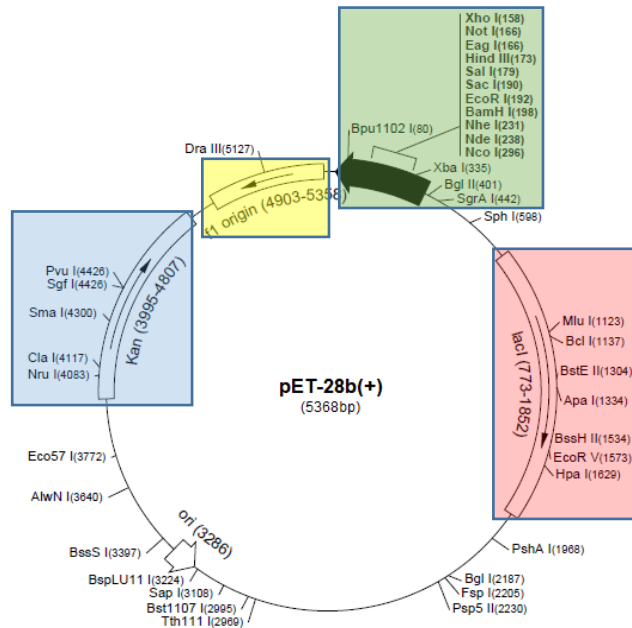
El DNA genómico de *P. aeruginosa* PAO1 se utilizó como molde para la reacción en cadena de la polimerasa (PCR) (Mullis *et al.*, 1986). Los siguientes oligonucleótidos sentido y antisentido se diseñaron con base en la secuencia de *pa2125*. Sentido: 5'-GGAATTCCATATGAACGGACACGCCAGACACTGG-3'

Antisentido: 5'-CCCAAGCTTTCAGGCCGCGCCGGTGGTCAG-3'

El oligonucleótido sentido contiene la secuencia de reconocimiento para la enzima de restricción *NdeI* (doble subrayado), un codón de iniciación (ATG, en negritas), una secuencia de anclaje para que la enzima de restricción se una mejor al sitio de reconocimiento (subrayado sencillo) y una secuencia de codificación para los siete primeros residuos de aminoácidos de la *PA2125*. El oligonucleótido antisentido contiene la secuencia de reconocimiento para la enzima de restricción *HindIII* (doble subrayado), una secuencia de anclaje de esta enzima (subrayado sencillo) y una secuencia de codificación para los siete últimos residuos de aminoácidos de la *PA2125*. Los oligonucleótidos se diseñaron con la ayuda de la herramienta en línea *OligoAnalyzer* 3.1 (*Integrated DNA Technologies*, <https://www.idtdna.com/calc/analyzer>) y se sintetizaron en la Unidad de Síntesis y Secuenciación de DNA (USSDNA) del Instituto de Biotecnología (IBT) de la UNAM, donde se utilizó el método del fosfito-triéster en fase sólida (McBride y Caruthers,

1983). La amplificación del gen se hizo usando la reacción en cadena de la polimerasa (PCR por sus siglas en inglés) en un equipo termociclador Techne modelo FTGENE2D (Cole-Parmer, RU), y una mezcla de DNA genómico (466 pmol/  $\mu$ L), los oligos sentido y antisentido (0.45 pmol/ $\mu$ L), MgCl<sub>2</sub> (2.0 mM), dNTPs (60  $\mu$ M), *pfu* amortiguador 1X (200 mM Tris-HCl (pH 8.8 at 25°C), 100 mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 100 mM KCl, 1% (v/v) Triton X-100, 1 mg/mL BSA) y DNA polimerasa *Pfu* (37.5 mU/  $\mu$ L). Para llevar a cabo la PCR, la mezcla se sometió a 95 °C durante 10 min, para la desnaturalización inicial del DNA genómico, y después se llevaron a cabo 30 ciclos, cada uno de los cuales consistió en 5 min de desnaturalización a 95 °C, 50 s de hibridación a 64 °C, 3 min de extensión (polimerización) a 72 °C y una extensión al final de los 30 ciclos durante 10 min a 72 °C.

El producto de PCR se analizó mediante electroforesis en agarosa y tras corroborar que tenía el tamaño esperado (1,449 pb) se purificó utilizando el *GeneJet PCR Purification Kit* (Thermo Scientific, EUA) y se digirió a 37 °C durante 16 h con las enzimas de restricción *NdeI* y *HindIII* (New England Biolabs, EUA) en el amortiguador *NEBuffer 2.1* (50 mM NaCl, 10mM Tris-HCl, 10mM MgCl<sub>2</sub>, 100 $\mu$ g/ml BSA y pH 7.9) (New England Biolabs, EUA). El gen digerido se clonó en el vector de expresión pET28b(+) (Novagen, Alemania) (Figura 5), que contiene una región de múltiples sitios de restricción en donde se hizo la inserción del gen *pa2125* usando las enzimas de restricción *NdeI* y *HindIII*. Este vector también posee un promotor y un terminador transcripcional para la RNA polimerasa T7, que produce altos niveles de transcripción en diferentes cepas de *E. coli* como la cepa BL21 Rosetta (DE3), así como un gen que confiere resistencia a kanamicina, que se usa como marcador de selección de las bacterias transformadas.



**Figura 6** Mapa del vector *pET-28b (+)* en el que se clonó el gen *pa2125*. La región sombreada en amarillo corresponde a la secuencia de nucleótidos del origen de la replicación/replicón (f1 ORI); en azul se marca la secuencia de nucleótidos del gen que confiere resistencia a kanamicina; en rojo, la secuencia de nucleótidos del gen que codifica la proteína represora lacI; en verde, la secuencia de nucleótidos de la región de clonación/expresión.

La ligación del vector pET28b(+) con el gen *pa2125* se realizó a una temperatura de 22 °C, por 16 h, usando *T4 DNA Ligase* (Thermo Scientific, EUA). El plásmido resultante se designó como pET28b-*pa2125*, el cual se usó para transformar células de *E. coli* de la cepa XL10-Gold por choque térmico. Las células transformadas se crecieron en placa de manera selectiva en presencia de kanamicina (concentración de 0.1 mg/mL en el medio de cultivo). De las colonias que crecieron se seleccionaron 4 al azar y se sembraron cada una por separado en medio líquido LB que contenía kanamicina (0,1 mg/mL) y se incubaron por 16 h a 37 °C con agitación constante a 200 rpm en una incubadora de agitación orbital MaxQ



4000 (Thermo Scientific, EUA). De estas células se extrajo plásmido pET28b-pa2125 utilizando *GeneElute Kit* (Sigma Aldrich, EUA). Como un método para comprobar la clonación del gen *pa2125* los plásmidos pET28b(+)-pa2125 extraídos de las 5 colonias que crecieron selectivamente se digirieron con las enzimas *NdeI* y *HindIII* y los productos respectivos se sometieron a electroforesis en geles de agarosa al 1.5 % (p/v). Cada gel se preparó por disolución de agarosa en amortiguador TAE (Tris 4 mM, acetato de sodio 8 mM, EDTA 100  $\mu$ M, a pH 7.5 ajustado con ácido acético glacial). Posteriormente, se adicionó bromuro de etidio 0.05  $\mu$ g/mL. Antes de la aplicación de las muestras en el gel, estas se mezclaron en una relación 1:10 con un amortiguador de carga que contenía azul de bromofenol 0.25 %, xilencianol 0.25 % y glicerol 50 % (v/v). La electroforesis se corrió en una cámara de electroforesis horizontal Mini-Sub<sup>®</sup> Cell GT (Bio-Rad, EUA) que contenía amortiguador TAE. Las bandas de DNA en los geles se observaron bajo luz UV. Una vez establecido cuales plásmidos contenían el inserto se procedió a juntarlos, la concentración final del plásmido pET28b-pa2125 fue 117 ng/ $\mu$ L. Para corroborar un adecuado marco de lectura y que no hubiera mutaciones en el gen, el plásmido se envió a secuenciar mediante el método de Sanger (Sanger y Coulson, 1975) en la Unidad de Síntesis y Secuenciación del IBT, UNAM. Para esto, se prepararon dos alícuotas de 16  $\mu$ L en tubos cónicos de 200  $\mu$ L, los cuales contenían plásmido a 50 ng/ $\mu$ L, adicional a esto a un tubo se adicionó oligonucleótido sentido del promotor T7 y a otro tubo se adicionó oligonucleótido antisentido del promotor T7.

#### **V.4 Sobreexpresión del gen *pa2125* y producción de la enzima recombinante *pa2125* en células de *E. coli***

Luego de comprobar que se contaba con el gen *pa2125* clonado en el vector de expresión, se transformaron con él células competentes de *E. coli* de la cepa BL21 (DE3) Rosetta (Novagen, Alemania) por medio de choque térmico. Para ello, primero se prepararon células competentes creciendo bacterias de la cepa BL21 Rosetta (DE3) en 10 mL de medio LB líquido durante 16 h a 37 °C y 200 rpm. De este cultivo se tomó una alícuota de 1.5 mL para inocular 50 mL de medio de LB y se incubó a 37 °C y 200 rpm hasta alcanzar un valor de densidad óptica a 600 nm ( $DO_{600}$ ) de 0.6 unidades de absorbancia. Posteriormente, el cultivo de células se mantuvo a 4 °C en un baño de hielo durante 20 min, tras lo cual se centrifugó a 6,000 rpm (4,637 g) por 10 min a 4 °C. El sobrenadante se decantó y la pastilla de células se resuspendió en 10 mL de una solución de  $CaCl_2$  100 mM, estéril y fría. La mezcla se colocó nuevamente en un baño de hielo durante 20 min y luego se centrifugó a 6,000 rpm (4,637 g) durante 10 min a 4 °C. El sobrenadante se desechó y la pastilla de células se resuspendió en 2 mL de cloruro de calcio 100 mM. Finalmente, se adicionó glicerol hasta obtener una concentración final de 20 % (v:v) y se hicieron alícuotas de 50  $\mu$ L que se almacenaron a -80 °C para su posterior uso. Para el paso de transformación, se tomó una alícuota de 50  $\mu$ L de las células competentes previamente preparadas y se añadieron 5  $\mu$ L del plásmido pET28b-*pa2125* (40 ng/ $\mu$ L); la mezcla se dejó en hielo durante 30 min. Luego de este tiempo, el tubo conteniendo la mezcla se pasó a un baño térmico a 42 °C durante 40 s e inmediatamente se pasó al hielo por 2 min más. Posteriormente, se agregaron 500  $\mu$ L de medio líquido LB estéril y la mezcla se incubó durante 1 h a 37 °C y 200 rpm. Transcurrido este tiempo, las células cultivadas se centrifugaron a 3,000 rpm por 2 min;  $\frac{3}{4}$  partes del sobrenadante se desecharon y con la

parte restante se resuspendió la pastilla de células que quedó. Con esta solución se sembraron placas con agar LB que contenían kanamicina a una concentración final de 0.1 mg/mL. Las placas se incubaron durante 16 h a 37 °C. De las colonias que crecieron se seleccionó una que se cultivó en medio LB líquido en presencia de kanamicina (0.1 mg/mL) a 37 °C y 200 rpm. De este cultivo se tomaron alicuotas de 700 µL que se pasaron a tubos estériles de 1.5 mL a los que se adicionaron 300 µL de glicerol estéril al 85% (v:v); las mezclas se homogenizaron mediante vórtex y se almacenaron a -80 °C para crioconservar la cepa transformada con el plásmido. Para producir la enzima recombinante, se inocularon 10 mL de LB líquido conteniendo kanamicina (0.1 mg/ml) con bacterias de la cepa de *E. coli* BL21 (DE3) Rosetta/pET28-pa2125 y se crecieron en agitación constante a 37 °C por 16 h; se tomó 1 mL de este cultivo para inocular 50 mL de LB líquido/kanamicina (0.1 mg/ml); las células se incubaron en agitación constante a 37 °C hasta alcanzar una densidad óptica a 600 nm ( $DO_{600}$ ) entre 0.5 y 0.7 unidades de absorbancia. En ese momento se indujo la síntesis de la enzima PA2125 con isopropil- $\beta$ -D-tiogalactopiranosido (IPTG) 0.1 mM, y se continuó la incubación por 4 horas más. Una vez terminado el tiempo de inducción con IPTG, el cultivo de células se centrifugó a 5,860 g durante 10 min se descartó el sobrenadante y la pastilla celular se almacenó a -70 °C para su posterior uso. El nivel de producción de la enzima se evaluó por electroforesis desnaturalizante en gel de poliacrilamida en presencia de SDS (SDS-PAGE), siguiendo el método de Laemmli (1970). Para ello, el cultivo de células transformadas se centrifugó a 5,860 g por 15 min. Se separó el sobrenadante del precipitado de células y a este último se le adicionó un volumen de amortiguador de fosfatos 50 mM, KCl 200 mM, etilén glicol 20% (v/v), EDTA 1 mM,  $\beta$ -mercaptoetanol 10 mM a pH 6.9; (amortiguador de extracción), igual al volumen de sobrenadante recuperado para resuspender la proteína contenida en este

precipitado. Posteriormente, en los carriles del gel de poliacrilamida se aplicó un volumen igual de la solución que contenía el precipitado que de la que contenía el sobrenadante. Por medio de la herramienta *ProtParam* del portal de recursos bioinformáticos *Expasy* (<http://web.expasy.org/protparam/>) se determinó la masa molar teórica de la enzima en 51.3 kDa, por lo cual se tomó positiva la presencia de la proteína cuando se observó una banda en una posición que corresponde a una masa molecular aproximada de 55 kDa. De esta forma se evalúa no sólo el nivel de producción de la enzima, sino también que proporción de esta está soluble y qué proporción agregada.

#### **V.5 Purificación a homogeneidad de la enzima PA2125**

El plásmido pET28b-pa2125 permitió producir la proteína PA2125 con una etiqueta de histidinas en el extremo N-terminal, lo que facilita su purificación en una columna de afinidad que contiene iones níquel (*Protino® Ni-NTA Agarose*; Macherey-Nagel, EUA), basándose en la interacción específica y reversible entre la etiqueta de histidinas y el níquel soportado en la resina. Para ello, las bacterias se crecieron en 4 L de LB líquido, como se describió anteriormente, tras lo cual se centrifugaron a 5,860xg durante 5 min, a 10 °C, desechando el sobrenadante. El precipitado de células se resuspendió en 160 mL de un amortiguador fosfatos 50 mM, KCl 200 mM, etilenglicol 20% (v/v), EDTA 1 mM,  $\beta$ -mercaptoetanol 10 mM a pH 6.9; (amortiguador de extracción). Luego se obtuvo el extracto crudo por ruptura de las células por ultrasonido en un sonicador VCX 500 *ultrasonic processor* (Sonics, EUA) a una amplitud de 23 % y en intervalos de 40 s encendido y 20 s apagado para completar un tiempo total de 15 min de sonicación. El lisado se clarificó mediante centrifugación a 26,900xg durante 20 min. El sobrenadante se pasó a través de la

columna de afinidad de 10 mL de volumen de cama, previamente equilibrada y lavada con 20 volúmenes de columna con el amortiguador de extracción. La concentración de proteína en el extracto crudo, sobrenadante, o el precipitado de células, así como en las fracciones obtenidas después de lavar la columna se cuantificaron por medio del método de Bradford (Bradford, 1976). La elución se realizó con 200 mL de una disolución de imidazol (Sigma-Aldrich, EUA) 150 mM disuelto en el amortiguador de extracción. Durante la elución se recogieron fracciones de 10 mL, a las que se le determinó espectrofotométricamente la concentración de proteína, utilizando el valor de absorptividad molar a 280 nm ( $50,420 \text{ M}^{-1} \text{ cm}^{-1}$ ) calculado en el portal de recursos bioinformáticos *Expasy* mediante la herramienta *ProtParam* (<http://web.expasy.org/protparam/>) (Gasteiger *et al.*, 2003). Esta herramienta usa el método descrito por Gill y von Hippel (1989), que se basa en la absorción a esta longitud de onda de los residuos aromáticos de la proteína. A estas fracciones se les midió su grado de pureza por SDS-PAGE y su actividad benzaldehído deshidrogenasa espectrofotométricamente a 30 °C, observando la aparición del producto NADH a  $\lambda = 340$  nm de acuerdo con el procedimiento descrito en Velasco-García *et al.* (2000). La reacción se inició por adición de la preparación de enzima a la mezcla de reacción que estaba constituida por un amortiguador de fosfato de potasio 100 mM a pH 8.0,  $\text{NAD}^+$  2 mM y benzaldehído 2 mM (ensayo estándar). Definimos una unidad de actividad enzimática como la cantidad de enzima que cataliza la formación de un  $\mu\text{mol}$  de NADH por minuto.

Debido a que la enzima PA2125 perdía actividad lentamente en presencia de imidazol, y dado el gran volumen de la preparación de enzima pura que hacía poco práctico eliminar el imidazol por filtración en gel e incluso por diálisis, se intentó eliminar el imidazol por sucesivos lavados en tubos concentradores Amicon, sin embargo el concentrar a la proteína durante dos lavados sucesivos cuando aún había imidazol presente llevaba a agregación

irreversible de la enzima. Por ello se decidió realizar una cromatografía de intercambio iónico para obtener una proteína libre de imidazol en un solo paso. Para esto se reunieron las fracciones eluidas de la columna de afinidad con imidazol 150 mM y que presentaron actividad benzaldehído deshidrogenasa y se diluyeron con un amortiguador de fosfatos 50 mM, etilenglicol 20% (v/v), EDTA 1 mM,  $\beta$ -mercaptoetanol 10 mM a pH 6.9 (amortiguador de dilución), hasta que se alcanzó una concentración de KCl 40 mM y de imidazol 30 mM. Luego se sometieron a una cromatografía de intercambio iónico usando una columna *Q Sepharose Fast Flow* (GE Healthcare, EUA) previamente equilibrada con un amortiguador de fosfatos 50 mM, KCl 40 mM, etilenglicol 20% (v/v), EDTA 1 mM,  $\beta$ -mercaptoetanol 10 mM a pH 6.9 (amortiguador de lavado). La columna se lavó usando 20 volúmenes de columna de este amortiguador y posteriormente la enzima se eluyó con el amortiguador de extracción, reuniéndose las fracciones del eluato que presentaron actividad benzaldehído deshidrogenasa. La concentración de proteína de esta preparación de enzima se determinó espectrofotométricamente, midiendo su absorción a  $\lambda = 280$  nm y usando el coeficiente de extinción molar 0.982 g/L ó 50,420 M<sup>-1</sup> cm<sup>-1</sup> calculado basándonos en la secuencia de aminoácidos de la proteína como se describió anteriormente. Finalmente, se hicieron alícuotas de 100  $\mu$ L de la preparación de enzima pura a una concentración de 1.0 mg/mL, y se almacenaron a -80 °C para su posterior uso en los estudios cinéticos.

#### **V.6 Determinación del estado oligomérico de la enzima PA2125**

El estado de asociación de la enzima se determinó por cromatografía de exclusión molecular (SEC por sus siglas en inglés) en una columna Superdex 200 HR-10/30 (Pharmacia Uppsala, Suecia) acoplada a un equipo de FPLC ÄKTA pure (GE Healthcare,

EUA). La columna se equilibró y se eluyó a temperatura ambiente con el amortiguador de extracción a una velocidad de flujo de 1.0 mL/min. Previamente a la cromatografía, la solución de proteína (500 µL) a una concentración de 10 µg/µL se filtró a través de una membrana hidrofílica de PVDF no estéril para jeringa con un poro de 0.22 µm de diámetro (Millipore, Millex-GV, Alemania). La elución se monitoreó siguiendo la absorbancia a 280 nm. La mezcla de proteínas estándar (Gel Filtration Standard, Bio-Rad, EUA) utilizada para la calibración de la columna contenía las proteínas: tiroglobulina de bovino, 670,000 Da, γ-globulina de bovino, 158,000 Da, ovoalbúmina de pollo, 44,000 Da, y mioglobina de caballo, 17,000 Da.

#### **V.7 Ensayos de actividad y determinación de los parámetros cinéticos para diferentes sustratos aldehídos**

Los ensayos de velocidad inicial ( $v_0$ ) se realizaron espectrofotométricamente a 30 °C, como se describió en Velasco-García *et al.* (2000). La reacción se inició con la adición de la enzima. Todos los ensayos enzimáticos se hicieron por duplicado, utilizando alícuotas de enzima provenientes de dos purificaciones diferentes. En los estudios de saturación por un aldehído a concentración fija del  $\text{NAD}^+$  o viceversa saturación por  $\text{NAD(P)}^+$  a concentración fija de un aldehído, los datos de velocidad inicial se ajustaron a la ecuación de Michaelis-Menten (Ecuación 6).

$$v_0 = V[S]/(K_m+[S]) \quad \text{(Ecuación 6),}$$

donde [S] es la concentración del sustrato variable,  $K_m$  es la concentración del sustrato a la que se tiene la mitad de la velocidad máxima,  $V$  es la velocidad máxima y  $K_{iS}$  es la constante de inhibición del sustrato variable.

En los estudios para obtener los patrones de velocidad inicial, los datos de velocidad inicial obtenidos se ajustaron de forma global a la ecuación correspondiente a un mecanismo Bi-Bi ordenado en estado estacionario (Ecuación 7):

$$v_0 = V [A][B]/(K_{ia}K_b + K_a[B] + K_b[A] + [A][B]) \quad (\text{Ecuación 7}),$$

donde  $v_0$  es la velocidad inicial determinada experimentalmente,  $V$  es la velocidad máxima estimada,  $[A]$  es la concentración de NAD(P)<sup>+</sup>,  $[B]$  es la concentración del aldehído,  $K_a$  y  $K_b$  son las constantes de Michaelis-Menten para NAD(P)<sup>+</sup> y para el aldehído, respectivamente, y  $K_{ia}$  es la constante de disociación del complejo E-NAD(P)<sup>+</sup>.

En todos los casos, para los ajustes de los datos de velocidad inicial se usó el programa de cómputo comercial Origin versión 5.0, formulado con el algoritmo de regresión no lineal de Marquardt (1963). Para efectos de visualización, con los parámetros cinéticos estimados con los ajustes a las respectivas ecuaciones, se construyeron gráficas de dobles recíprocos ( $1/v_0$  frente a  $1/[S]$ ).

## V.8 Ensayos de cinética rápida

Se midió el curso temporal de las reacciones NAD<sup>+</sup>-dependientes de la oxidación del benzaldehído y tres de sus derivados (4-metoxibenzaldehído, 3-nitrobenzaldehído y 4-nitrobenzaldehído) catalizadas por la enzima PA2125 siguiendo la formación del nucleótido reducido NADH en un intervalo de milisegundos, usando concentraciones saturantes de sustratos. Para estos estudios se utilizó un equipo de flujo detenido (SX-18MV de Applied Photophysics®). La aparición del NADH se siguió por la absorbancia a 340 nm en un detector de luz UV/Vis. En el caso en que los cursos temporales mostraron un estallido (“burst”) previo al establecimiento de la velocidad de estado estacionario, los datos se



ajustaron a una ecuación que incluye la de una exponencial sencilla creciente y la de una recta (Ecuación 8):

$$[\text{NADH}] = A_{burst} (1 - e^{-k_{burst}t}) + v_{ss}t \quad (\text{Ecuación 8})$$

donde  $A_{burst}$  es la amplitud del estallido observado,  $k_{burst}$  es la constante de velocidad de orden uno de este estallido (en unidades de  $s^{-1}$ ),  $v_{ss}$  es la velocidad de estado estacionario (pendiente de la fase lineal) y  $t$  es el tiempo.

### **V.9 Obtención de un modelo por homología de la estructura tridimensional de la enzima PA2125**

El modelo por homología obtenido para la enzima PA2125 se construyó con la colaboración del Dr. Rogelio Rodríguez Sotres, Departamento de Bioquímica, Facultad de Química, UNAM. La búsqueda de la estructura cristalina de ALDH más apropiada para ser usada como molde se llevó a cabo con Turingen's suit (Söding *et al.*, 2005), usando tanto el programa HHpred (Söding, 2005) como el programa HHsensor (Söding *et al.*, 2006). Se utilizó el método Rd.HMM (Martínez-Castilla y Rodríguez-Sotres, 2010) para medir la habilidad (“*fitness*”) de estos posibles moldes para albergar la secuencia de aminoácidos de PA2125. Con la secuencia de aminoácidos del molde seleccionado se hicieron alineamientos con la secuencia de aminoácidos de PA2125 usando el servidor HHpred, así como un alineamiento manualmente curado tomando la alineación Rd.HMM en el archivo de búsqueda HMMer como referencia. Se construyeron modelos por homología de una subunidad con el programa MODELLER 9v2 (Eswar *et al.*, 2008), cada uno de ellos se sometió a optimización de geometría usando el campo de fuerza AMBER99SB implementado en el programa informático Chimera (Pettersen *et al.*, 2004) y se relajaron

con ROSETTA suite (Raman *et al.*, 2009). Se usó el modelo de subunidad mejor calificado con Rd.HMM para ensamblar un tetrámero con la ayuda de los programas computacionales *Visual Molecular Dynamics* (VMD) (Humphrey *et al.*, 1996) y MultiSeq (Roberts *et al.*, 2006). Las incoherencias estructurales en el tetrámero resultante se corrigieron utilizando el protocolo ROSETTA *suit fast: relax* (Raman *et al.*, 2009). El modelo del tetrámero resultante se mantuvo en una caja octaédrica periódica con una distancia de 1.2 nm a la pared de la caja, para generar un sistema periódico infinito, y se sometió a simulaciones de dinámica molecular (MD) de 15 ns en solvente explícito con 0.15 M NaCl formando un sistema neutro en el campo de fuerza AMBER99SB, utilizando GROMAX 4.6.7 con un esquema NPT, con *Particle-Mesh-Ewald* (PME *electrostatics*) para el cálculo de la electrostática y con LINCS para restringir las vibraciones de los enlaces C-H. Al final de la simulación, la estructura modelada se sometió a una minimización de energía por gradiente conjugado, en su misma caja de solvente e iones y usando el mismo campo de fuerzas. La estructura minimizada se extrajo y se analizó por Rd.HMM (Martínez-Castilla y Rodríguez-Sotres, 2010) para validar su calidad y consistencia estructural.

#### **V.10 Acoplamiento molecular de los sustratos en el sitio activo de PA2125**

Para investigar las bases estructurales de la especificidad de la enzima PA2125 por los sustratos probados en los ensayos de cinética enzimática, se llevaron a cabo acoplamientos moleculares (“docking”) entre estos aldehídos y el sitio activo del modelo por homología de la estructura tridimensional de PA2125. Para ello se utilizó el programa AutoDock Vina (Trott y Olson 2010) del software Chimera (<http://www.rbvi.ucsf.edu/chimera>) usando el modelo por homología de PA2125 como receptor, el cual se mantuvo rígido durante el

acoplamiento molecular. Los ligandos utilizados contenían los átomos pesados (C, N, O) y los hidrógenos (H). Todas las estructuras de los aldehídos se convirtieron al formato PDB, para lo cual se construyeron utilizando el programa Spartan '14 v.1.1.8 ([https://www.wavefun.com/products/windows/Student/win\\_student.html](https://www.wavefun.com/products/windows/Student/win_student.html)), teniendo cuidado de poner las cargas cuando era necesario. Con la ayuda de este mismo software se hizo una optimización tridimensional de estas estructuras realizando cálculos de equilibrio de conformeros a través de cálculos semiempíricos con el campo de fuerza PM6 (Řezáč *et al.*, 2009). Finalmente los archivos con las estructuras se almacenaron en formato PDB. Se obtuvieron las coordenadas del átomo de azufre de la cisteína catalítica Cys282 del modelo por homología de PA2125, (61.90, 80.16, 21.62), que correspondían al centro de la caja 8x8x8 Å que cubre el sitio de unión seleccionado para los ligandos, se obtuvieron mediante el programa computacional para visualización molecular Coot (Emsley & Cowtan, 2004), con la herramienta *Go to atom*. Posteriormente, se ingresaron estas coordenadas en el modelo previamente generado de la enzima, usando la opción *Surface/Binding Analysis* de la pestaña *Tools* del software Chimera. Las opciones de optimización para el receptor y para el ligando dadas por el programa no se modificaron. Finalmente se analizaron los 10 modos de unión generados por el programa y se seleccionó aquella que mostraba mayor concordancia con lo reportado en la literatura sobre el mecanismo químico de las ALDHs (revisado en Muñoz-Clares *et al.*, 2010; Muñoz-Clares *et al.*, 2011), es decir, se usaron como criterios el que el oxígeno del aldehído estuviera a distancia de puente de hidrógeno del nitrógeno amídico de la cadena principal de la Cys282 y del nitrógeno amídico de la cadena lateral de la Asn149, y el carbono a una distancia razonable del grupo tiol de la cisteína catalítica para que pudiera llevarse a cabo el ataque nucleofílico. Una vez obtenidas

los mejores modos de unión para cada aldehído se hicieron imágenes usando el programa Pymol (<https://pymol.org/2/>).

## VI RESULTADOS

### VI.1 Ortólogos de PA2125

Con el fin de investigar si la enzima PA2125 poseía algún ortólogo en la cepa PAO1 de *P. aeruginosa*, se buscaron todos los genes que codificaban para enzimas pertenecientes a la superfamilia ALDH en el genoma secuenciado de esta cepa. Se encontraron 23 miembros de esta superfamilia (Tabla 1), es decir tres menos de los reportados por Sophos y Vasiliou (2003). Estas tres enzimas son: una enzima referida por estos autores como N/A con código de acceso a la base de datos UniProtKB Q9APW7 (acceso NCBI: WP\_023091335) que encontramos que tiene asignado el cd07092 el cual corresponde a la familia ALDH26 cuyos miembros son  $\gamma$ -aminobutiraldehído deshidrogenasas y que a pesar de que se encuentra en otras especies del género *Pseudomonas* y en algunas cepas de *P. aeruginosa* no lo está en la cepa PAO1; la enzima codificada por el gen *exaC* con código de acceso a la base de datos UniProtKB Q9ZAA1, que duplica otra ya considerada por estos autores; y la enzima codificada por el gen *pahF* (código de acceso al NCBI: BAA12243.1) que encontramos que tiene asignado el cd07105 con función de salicilaldehído deshidrogenasa y que está presente en otras cepas de *P. aeruginosa* pero no en PAO1. Analizamos la secuencia de las 23 ALDHs encontradas para poderlas asignar a una familia ALDH ya numerada, o a un determinado grupo cd cuando aún no se les ha asignado una familia dentro de la superfamilia de las ALDHs, encontrando que 3 pertenecen a la familia ALDH6

(cd07085) cuyos miembros se sabe que son ALDHs acilantes, 3 pertenecen a la familia ALDH27 (cd07112), de las cuales una de ellas ya ha sido caracterizada como aminoaldehído deshidrogenasa, 2 están anotadas como  $\alpha$ -cetoglutarato semialdehído deshidrogenasas (cd07129), una anotada como N-Succinilglutamato 5-semialdehído deshidrogenasa (cd07095), una anotada como prolina deshidrogenasa (cd07125), dos anotadas solamente como ALDH, pero que encontramos que pertenecen a la misma familia puesto que tienen asignado el mismo número cd (cd07559), una enzima perteneciente a cada una de las siguientes familias: ALDH5 (cd07103), ALDH7 (cd07130), ALDH8 (cd07093), ALDH9 (cd07090) y ALDH11 (cd07082), y 6 que pertenecen a familias que contienen un solo representante en *P. aeruginosa* PAO1 y que aunque aún no han sido numeradas consideramos que son familias ALDHs atendiendo a que poseen diferentes números cd. Entre estas últimas se encuentra la PA2125. Hay que mencionar que las anotaciones en las bases de datos que indican una actividad fueron hechas en base a homología con enzimas caracterizadas en otras bacterias, pero que a la fecha en *P. aeruginosa* sólo se han caracterizado la enzima ALDH9 (la PaBADH, PA5373), una enzima de la familia ALDH27 (la PA5312) y se ha reportado la estructura cristalográfica de la PA0898 (código PDB 3JU8).

**Tabla 1. Enzimas pertenecientes a la superfamilia ALDH en *P. aeruginosa* cepa PA01.**

Nombre del gen	Locus	Anotación en el genoma ( <a href="http://www.pseudomonas.com/">http://www.pseudomonas.com/</a> )	Acceso NCBI	CDD NCBI (Familia)	Actividad según dominio conservado (CDD NCBI)	Función fisiológica (referencia)	Identidad respecto a PA2125 (%)
<i>gapN</i> , <i>gapB</i>	PA2323	Probable gliceraldehído-3-fosfato deshidrogenasa	NP_251013	cd07082 (ALDH11)	Gliceraldehído-3-fosfato deshidrogenasa no fosforilante	Generación de NADPH (probable)	27.4
<i>bauC</i>	PA0130	Malonato semialdehído deshidrogenasa	NP_248820	cd07085 (ALDH6)	Metilmalonato semialdehído deshidrogenasa	Degradación de poliaminas (Yao <i>et al.</i> , 2011)	29.1
<i>pa0747</i>	PA0747	Probable aldehído deshidrogenasa	NP_249438	cd07085 (ALDH6)	Metilmalonato semialdehído deshidrogenasa	Degradación de valina (probable)	30.3
<i>mmsA</i>	PA3570	Metilmalonato semialdehído deshidrogenasa	NP_252260	cd07085 (ALDH6)	Metilmalonato semialdehído deshidrogenasa	Desconocida	27.7
<i>betB</i>	PA5373	Betaína aldehído deshidrogenasa	NP_254060	cd07090 (ALDH9)	Betaína aldehído deshidrogenasa	Degradación de colina y síntesis de glicina betaína (Nagasawa, <i>et al.</i> , 1976)	33.8
<i>hpcC</i> , <i>hpaE</i>	PA4123	5-Carboxi-2-hidroxiuconato semialdehído deshidrogenasa	NP_252812	cd07093 (ALDH8)	2-Hidroxiuconato semialdehído deshidrogenasa	Degradación de compuestos aromáticos (probable)	34.4
<i>aruD</i> , <i>astD</i>	PA0898	N-Succinilglutamato 5-semialdehído deshidrogenasa	NP_249589	cd07095 (No tiene familia asignada)	N-Succinilglutamato 5-semialdehído deshidrogenasa	Degradación de arginina (Itoh, 1997; Jann <i>et al.</i> , 1986)	31.7
<i>davD</i> , <i>gabD</i>	PA0265	Glutárico semialdehído deshidrogenasa	NP_248956	cd07103 (ALDH5)	Succinato semialdehído deshidrogenasa	Degradación de poliaminas (Yao <i>et al.</i> , 2011)	35.6
<i>pa0219</i>	PA0219	Aldehído deshidrogenasa	NP_248910	cd07112 (ALDH27)	$\gamma$ -Glutamil- $\gamma$ -aminobutiraldehído deshidrogenasa	Desconocida	34.1
<i>pa4189</i>	PA4189	Probable aldehído deshidrogenasa	NP_252878	cd07112 (ALDH27)	$\gamma$ -Glutamil- $\gamma$ -aminobutiraldehído deshidrogenasa	Desconocida	34.3
<i>kauB</i> , <i>pauC</i>	PA5312	Trimetil amonibutiraldehído deshidrogenasa	NP_253999	cd07112 (ALDH27)	$\gamma$ -Glutamil- $\gamma$ -aminobutiraldehído deshidrogenasa	Degradación de arginina y poliaminas (Jann <i>et al.</i> 1998; Yao <i>et al.</i> , 2011; Chou <i>et al.</i> , 2013)	32.1
<i>pa3504</i>	PA3504	Probable aldehído deshidrogenasa	NP_252194	cd07114 (No tiene familia asignada)	Aldehído deshidrogenasa no caracterizada	Desconocida	37.0

<i>pa2125</i>	PA2125	Aldehído deshidrogenasa	NP_250815	cd07120 (No tiene familia asignada)	Furfural deshidrogenasa	Metabolismo de ácido furóico (Nichols y Mertens, 2013)	100
<i>putA</i> <i>pruA</i>	PA0782	Prolina deshidrogenasa	NP_249473	cd07125 (No tiene familia asignada)	Glutamato semialdehído deshidrogenasa	Degradación de prolina (Nakada <i>et al.</i> , 2002)	28.8
<i>lhpG</i>	PA1253	$\alpha$ -Cetoglutarato semialdehído deshidrogenasa	NP_249944	cd07129 (No tiene familia asignada)	$\alpha$ -Cetoglutarato semialdehído deshidrogenasa	Degradación de 4-hidroxi-L-prolina (Li y Lu, 2016)	28.9
<i>pa2217</i>	PA2217	Probable aldehído deshidrogenasa	NP_250907	cd07129 (No tiene familia asignada)	$\alpha$ -Cetoglutarato semialdehído deshidrogenasa	Degradación de 4-hidroxi-L-prolina	25.6
<i>amaB</i> , <i>pcd</i>	PA1027	$\Delta$ -1-Piperidina-6-carboxilato deshidrogenasa	NP_249718	cd07130 (ALDH7)	$\alpha$ -Amino adipato semialdehído deshidrogenasa	Degradación de L-lisina (probable) (Revelles <i>et al.</i> , 2005)	28.4
<i>pa0366</i>	PA0366	Probable aldehído deshidrogenasa	NP_249057	cd07133 (No tiene familia asignada)	Coniferil aldehído deshidrogenasa	Degradación de eugenol (Achterholt <i>et al.</i> , 1998)	28.7
<i>pa4899</i>	PA4899	Probable aldehído deshidrogenasa	NP_253586	cd07151 (No tiene familia asignada)	p-hidroxibenzaldehído deshidrogenasa	Degradación de p-cresol Probable	33.7
<i>exaC</i>	PA1984	Aldehído deshidrogenasa dependiente de NAD <sup>+</sup> ExaC	NP_250674	cd07559 (No tiene familia asignada)	Acetaldehído deshidrogenasa II	Metabolismo de etanol (Schobert y Görisch, 1999)	31.9
<i>hdhA</i> , <i>exaC2</i>	PA4022	Hidrazona deshidrogenasa	NP_252711	cd07559 (No tiene familia asignada)	Acetaldehído deshidrogenasa II	Degradación de ácido adípico (Taniyama <i>et al.</i> , 2012)	31.4
<i>pa4073</i>	PA4073	Probable aldehído deshidrogenasa	NP_252762	No asignado (No tiene familia asignada)	Aldehído deshidrogenasa	Desconocida	36.9
<i>proA</i>	PA4007	$\gamma$ -Glutamil fosfato reductasa	NP_252696	No asignado (No tiene familia asignada)	$\gamma$ -Glutamil fosfato reductasa	Biosíntesis de prolina	

Mediante alineamientos múltiples de las secuencias de aminoácidos de estas ALDHs se encontró que ninguna posee identidad  $\geq 40\%$  con la PA2125 y que por tanto pertenecen a familias diferentes, consistente con que tienen asignados diferentes número cd. Este

resultado, por tanto, no permitió inferir una posible función fisiológica de la enzima PA2125, ni siquiera conocer el tipo de aldehídos que podría usar esta enzima como sustratos, por lo que extendimos la búsqueda de ortólogos usando las secuencias de enzimas ALDHs reportadas a la fecha en la base de datos NR (no redundantes) del NCBI, y consideramos sólo aquellas secuencias cuya identidad con PA2125 fue  $\geq$  al 42 % para encontrar las que pertenecen a la misma familia. Esto se corroboró viendo que tenían asignado el número cd07120. Hicimos el corte en 42 % de identidad porque aquellas que poseían identidad  $\leq$  42 % ya tenían asignado otro número cd.

Encontramos un total de 179 secuencias (Anexo 1), todas ellas bacterianas, y que habían sido anotadas como betaína aldehído deshidrogenasa (5), 2,5-dihidroxibenzaldehído (gentisalaldehído) deshidrogenasa (1), furfural deshidrogenasa (1), succinato semialdehído deshidrogenasa (1), PutA o glutamato semialdehído deshidrogenasa no fosforilante (1), acil-CoA reductasa (2) o, la mayor parte de ellas, sólo como ALDHs de (168). En la Tabla 2 se incluyen ejemplos de estas enzimas ALDHs que han sido anotadas con alguna actividad específica y que presentan una identidad mayor al 47 % con PA2125.

A primera vista varias de estas anotaciones son incorrectas. Una de ellas es la anotación de acil-CoA reductasa, puesto que estas proteínas son reductasas de ácidos grasos responsables de la síntesis de aldehídos de cadena larga que son sustratos en la reacción luminiscente catalizada por la luciferasa y que tienen asignado el cd07080. De hecho, un alineamiento de una de las enzimas de la familia de la PA2125 así anotadas con una acil-CoA reductasa *bona fide*, la de *Photobacterium kishitanii*, da un 0% de identidad.



**Tabla 2. Porcentaje de identidad con PA2125 de enzimas seleccionadas de su misma familia y otras enzimas ALDHs que poseen la actividad asignada a algunas de las enzimas de la familia de la PA2125.**

Acceso NCBI	Descripción	Identidad con respecto PA2125 (cd07120) (%)	Identidad con respecto PaBADH PA5373 (cd07090) (%)	Identidad con respecto Gentisaldehído <i>Burkholderia</i> sp. BC1 (cd07120) (%)	Identidad con respecto psfA de <i>P. putida</i> (cd07120) (%)	Identidad con Succinato semialdehído deshidrogenasa <i>P. aeruginosa</i> (PRK11241) (%)	Identidad con respecto PutA de <i>P. putida</i> (PRK11809) (%)	Identidad con respecto acil-CoA reductasa [ <i>Photobacterium kishitanii</i> ] (cd07080) (%)
NP_250815.1	Aldehído deshidrogenasa PA2125 [ <i>Pseudomonas aeruginosa</i> ] (cd07120)	100	36	60	47	35	29	21
SCY80161.1	Betaína aldehído deshidrogenasa [ <i>Acinetobacter baumannii</i> ] (cd07120)	99	36	60	47	35	29	21
NP_254060.1	Betaína aldehído deshidrogenasa PA5373 [ <i>Pseudomonas aeruginosa</i> ] (cd07090)	36	100	33	33	36	22	25
KTK92484.1	Aldehído deshidrogenasa [ <i>Pseudomonas putida</i> ] (cd07120)	99	37	60	47	35	30	27
APP18125.1	Gentisaldehído deshidrogenasa [ <i>Burkholderia</i> sp. BC1] (cd07120)	60	33	100	43	33	30	35
ACA09737.1	PsfA [ <i>Pseudomonas putida</i> ] (cd07120)	47	33	43	100	35	29	24
AIJ45734.1	succinate-semialdehyde dehydrogenase [ <i>Comamonas testosteroni</i> TK102] (cd07120)	45	33	42	64	36	28	22

AAG03654.1	Succinato semialdehído deshidrogenasa [ <i>P. aeruginosa</i> ] (CDD NCBI: PRK11241)	36	38	35	36	100	31	24
ALP00403.1	PutA [ <i>Pseudomonas putida</i> ] (cd07120)	42	32	43	48	29	27	24
AAF25000.1	PutA [ <i>Pseudomonas putida</i> ] (CDD NCBI: PRK11809)	29	27	30	29	31	100	23
SDG68883.1	Acyl-CoA reductase [ <i>Streptomyces jietaisiensis</i> ] (cd07120)	48	34	50	45	33	29	0.0
WP_045042154	acyl-CoA reductase [ <i>Photobacterium kishitanii</i> ] (cd07080)	0.0	0.0	0.0	0.0	0.0	23	100

Por tanto, concluimos que estas enzimas sí son de la misma familia de la PA2125 (cd07120) pero con función aún desconocida. Igualmente, la enzima de la familia cd07120 anotada como succinato semialdehído deshidrogenasa tampoco creemos que posee esta actividad, puesto que su identidad con las succinato aldehído deshidrogenasas (cuyo cd es 07103 y pertenecen a la familia ALDH5) es menor al 42 %; de hecho, su identidad con la enzima PA0265 de *P. aeruginosa* que posee esta actividad es del 36 %. También está equivocada la anotación de putA, como lo indica la identidad muy por debajo del 40% de la enzima así asignada con verdaderas enzimas putA (con código CDD NCBI: PRK11904), de las cuales *P. aeruginosa* posee una enzima la PA0782 con una identidad del 28.8 % con PA2125, y el hecho de que estas enzimas son muy diferentes porque son bifuncionales, es decir poseen dos dominios uno con actividad de prolina deshidrogenasa y otro con

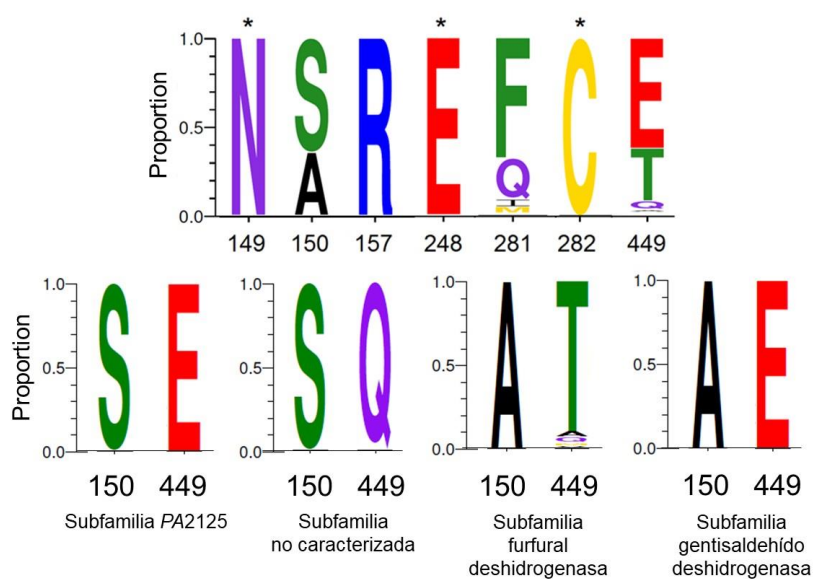
actividad de glutamato semialdehído deshidrogenasa no-fosforilante catalizando la formación de glutamato a partir de glutamato semialdehído. Finalmente, no creemos que las enzimas de la familia de la PA2125 anotadas como betaína aldehído deshidrogenasa lo sean, de nuevo basándonos en su bajo grado de identidad con las betaína aldehído deshidrogenasas de  $\gamma$ -proteobacterias (cd07090; ALDH9), de las que existe una enzima en *P. aeruginosa* PAO1, la PA5373 (*PaBADH*), que ha sido extensivamente caracterizada bioquímica y estructuralmente por nuestro grupo de investigación (Velasco-García *et al.*, 1999; Velasco-García *et al.*, 2000; González-Segura, *et al.*, 2002; Velasco-García *et al.*, 2003; Velasco-García *et al.*, 2006; González-Segura, *et al.*, 2009; Zaldívar-Machorro, *et al.*, 2011; Velasco-García *et al.*, 2013). En cuanto a las anotaciones de una enzima como 2,5 dihidroxibenzaldehído (gentisalaldehído) deshidrogenasa y de otra como furfural deshidrogenasa sí son ciertas, puesto que en ambos casos se ha comprobado experimentalmente que poseen estas actividades. La actividad de la primera se propuso en un estudio genómico (Chowdhury *et al.*, 2014) en el que se observó su participación en la ruta de degradación del ácido 2-hidroxi-1-naftoico presente en aguas residuales oxidando gentisalaldehído a gentisato. La actividad furfural deshidrogenasa se asignó a la enzima codificada por el gen *psfA* de *P. putida* (acceso NCBI: ACA09736.1) porque la transcripción de este gen se indujo cuando la bacteria se creció en alcohol furfúrico, furfural o ácido furoico (Nichols y Mertens, 2008), compuestos que se producen en suelo por la degradación de material vegetal lignocelulósico. *P. aeruginosa* PAO1 no posee ninguna enzima anotada con estas actividades, por lo que estos aldehídos, o aldehídos aromáticos relacionados, sí pudieran ser sustratos de la PA2125. La identidad de la enzima gentisalaldehído deshidrogenasa con PA2125 es relativamente alta (60%), lo que apoya aún

más esta idea. Así mismo, la identidad de la enzima de *P. putida* a la que se le asignó la actividad furfural deshidrogenasa con PA2125 es del 47 %, indicando que ambas enzimas son de la misma familia.

Para investigar si la PA2125 pudiera tener alguna de estas dos actividades asignadas a dos enzimas de su misma familia seguimos dos estrategias: 1) Comparar los residuos del sitio activo que pudieran estar determinando la especificidad por el aldehído de la PA2125 con los de las otras dos enzimas de su misma familia cuya actividad ya ha sido reportada. 2) Analizar y comparar los contextos genómicos de estas enzimas y de la PA2125 (y para confirmar que la anotación como betaína aldehído deshidrogenasa es un error, examinamos también el contexto genómico de la PaBADH). El comparar los contextos genómicos nos ayuda a determinar si la enzima que nos interesa participa en alguna ruta metabólica ya caracterizada y que forma parte de algún operón ya estudiado. Así mismo, podemos inferir cuál podría ser el compuesto que al transformarse produce el aldehído sustrato de la ALDH bajo estudio, en caso de que la enzima que cataliza esa reacción ya está caracterizada o pertenezca a alguna familia de enzimas con un cd asignado. 3) Determinar la actividad de la PA2125 con los aldehídos sugeridos como sustratos para estas enzimas y con aldehídos relacionados.

1) *Residuos del sitio activo posibles determinantes de la especificidad para el aldehído:* Los miembros de la familia ALDH con cd07120, de la cual forma parte PA2125, tienen conservados los residuos catalíticos asparagina, glutámico y cisteína (N149, E248 y C282, numeración correspondiente a PA2125) lo cual indica que son ALDHs hidrolíticas. En cuanto a los residuos que pueden participar en la especificidad por el sustrato en esta familia, encontramos que en la posición 150 de la PA2125 puede haber una serina o una alanina, indicando que en esta posición debe existir un residuo pequeño, en la posición 157

de la PA2125 siempre hay una arginina, lo cual sólo se ha encontrado hasta la fecha en las enzimas de la familia cd07120, y en la posición 449 puede haber un glutámico o treonina. Atendiendo al criterio del tipo de residuos que existen en esta posición, parecen existir al menos cuatro subfamilias: una, a la que pertenece la PA2125, cuyos miembros poseen Ser150 y Glu449 (numeración de PA2125), otra subfamilia cuyos miembros tienen secuencias con serina y glutamina es las posiciones equivalentes, otra a la que pertenece la gentaldehído deshidrogenasa, cuyos miembros poseen alanina y glutámico en las posiciones equivalentes, y otra a la que pertenece la furfural deshidrogenasa, que contiene secuencias que en estas posiciones poseen alanina y treonina, respectivamente (Figura 7).



**Figura 7. Logos de los residuos conservados de la familia ALDH cd07120.** (A) Logos donde se muestran marcados con asterisco los residuos catalíticos en las posiciones 149, 248, 282, y los residuos que participan en la especificidad por sustrato en las posiciones 150, 157, 281 y 449 de acuerdo con la numeración en PA2125. (B) Logos donde se muestran las posibles cuatro subfamilias teniendo en cuenta los residuos en las posiciones 150 y 449. Se obtuvieron a partir de alineamientos múltiples de PA2125 y enzimas ortólogos con la herramienta BLAST del NCBI (<https://blast.ncbi.nlm.nih.gov/blast.cgi>).

Los resultados de este análisis nos sugieren que el residuo en la posición correspondiente al Glu449 de PA2125 es importante para determinar la especificidad dentro esta familia y que probablemente la PA2125 no sería una furfural deshidrogenasa, pero sí podría ser una gentisaldehído deshidrogenasa u oxidar a un sustrato químicamente parecido al gentisaldehído.

2) *Contextos genómicos.* Los contextos genómicos de la PA2125 y de aquellas enzimas de su misma familia que están correctamente anotadas, la gentisaldehído deshidrogenasa de *Burkholderia sp.* BC1 y la furfural deshidrogenasa de *P. putida*, así como el de la PaBADH se muestran en la Figura 8.



*Pseudomonas aeruginosa* PAO1

PA2122: Policétido ciclasa/deshidratasa; superfamilia cl14643; familia pfam10604; PA2123: regul. transc. LysR; superfamilia cl28285; familia cd08422; PA2124: deshidrogenasa; superfamilia cl27052; familia pfam00732; **PA2125: ALDEHÍDO DESHIDROGENASA**; familia cd07120.



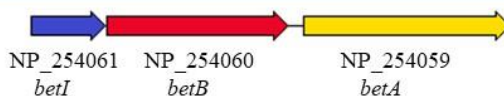
*Pseudomonas putida* Fu1

ACA09736: alcohol deshidrogenasa; superfamilia cl27753; **ACA09737: FURFURAL DESHIDROGENASA**; familia cd07120; ACA09738: regul. Transcrip. LysR; superfamilia cl28285; familia COG0583; ACA09739 hidroxibenzoato carboxiliasa; superfamilia cl19190; familia PRK06029.



*Burkholderia sp.* BC1

APP18123: regul. Transc. TetR\_N; superfamilia cl27689; familia COG1309; APP18124: dihidrodiol deshidrogenasa; superfamilia cl27754; **APP18125: GENTISALDEHÍDO DESHIDROGENASA**; familia cd07120; APP18126: 2-naftol dioxigenasa; superfamilia cl00938; familia cd08881; APP18127: 2-naftol dioxigenasa; superfamilia cl09109; APP18128: ferredoxina; superfamilia cl00159; APP18129: ferredoxin reductasa; superfamilia cl26176; familia pfam14759.



*Pseudomonas aeruginosa* PAO1

NP\_254061: regul. Transc. LysR; superfamilia cl27691; familia PRK00767; **NP\_254060: BETAÍNA ALDEHÍDO DESHIDROGENASA**; familia cd07090; NP\_254059: colina deshidrogenasa; superfamilia cl27052; familia PRK02106.

**Figura 8. Contextos genómicos de los genes que codifican para PA2125, las enzimas de su misma familia con actividad probada y la PaBADH**

Al observar los contextos genómicos de las anteriores enzimas encontramos que en todos ellos el gen que antecede al que codifica para la aldehído deshidrogenasa es el que codifica para la enzima que cataliza la reacción en la que se produce el aldehído. En el contexto genómico de PA2125 esta enzima es una deshidrogenasa (PA2124) que pertenece a la superfamilia de las glucosa-metanol-colina (GMC) oxidorreductasas (con código CDD NCBI cl27052). En el de la furfural deshidrogenasa codificada por el gen *psfA* de *P. putida* es una alcohol deshidrogenasa codificada por el gen *psfG* (acceso NCBI: 09736.1) de la superfamilia de alcoholes deshidrogenasas de cadena corta (con código CDD NCBI cl27753). De nuevo, los resultados de este análisis nos sugieren que la PA2125 no es una furfural deshidrogenasa, lo que está además apoyado por nuestro hallazgo de que *P. putida* posee otra enzima no caracterizada que tiene un grado de identidad del 99 % con la PA2125 y cuyo gen se encuentra en un contexto genómico similar al del gen *pa2125*. Es altamente improbable que una bacteria tenga dos enzimas con igual actividad fisiológica bajo dos contextos genómicos diferentes. En el contexto genómico de la gentisaldehído deshidrogenasa de *Burkholderia sp.* BC1 hay una hidratasa aldolasa codificada por el gen *nolE* (acceso NCBI: APP18130.1) la cual pertenece a la superfamilia aldolasa clase I con código CDD NCBI: cl28880. Pareciera entonces que de acuerdo a este criterio la PA2125 tampoco poseería la actividad de gentisaldehído deshidrogenasa. Por último, en el contexto genómico de la PaBADH (PA5373, con código acceso NCBI: NP\_254060.1), la enzima encargada de producir el aldehído es la colina deshidrogenasa (PA5372, con código de acceso al NCBI: NP\_254059.1) codificada por el gen *betA* (código CDD NCBI:

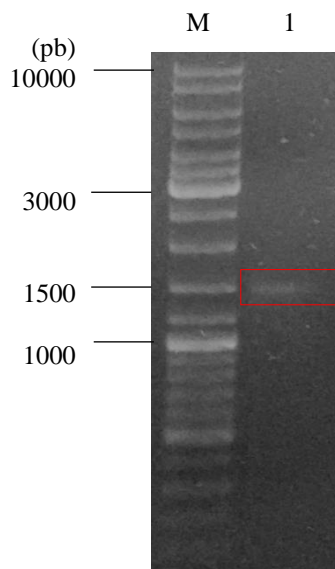
PRK02106) y que pertenece a la familia de enzimas de la superfamilia de GMC oxidorreductasas al igual que la PA2124 cuyo gen que se encuentra en el contexto genómico del gen de la PA2125. Probablemente por esta razón varias enzimas de las pertenecientes a la familia de la PA2125 se han anotado como betaína aldehído deshidrogenasas. Además, en el contexto del gen de la PA2125, al igual que en los otros contextos genómicos aquí analizados, existe un gen que codifica para un regulador transcripcional. En el operón de PA2125 es el que codifica para la proteína PA2123, que pertenece a la superfamilia de reguladores transcripcionales LysR (CDD NCBI: c128285) y se le ha asignado el cd08422. El gen del regulador transcripcional en el contexto genómico del gen de la furfural deshidrogenasa de *P. putida* también codifica a una proteína que pertenece a la superfamilia LysR pero es de la familia COG0583. En el contexto genómico de la gentisaldehído deshidrogenasa de *Burkholderia sp.* el regulador transcripcional pertenece a la familia COG1309 de la superfamilia TetR\_N (CDD NCBI: c127689). Por último el gen del regulador transcripcional que forma parte del operón de la PaBADH codifica una proteína que pertenece a la superfamilia de reguladores transcripcionales TetR con CDD NCBI: c127691 y dentro de ésta a la familia PRK00767, conocida como la familia de reguladores transcripcionales BetI. Esto nos indica que el regulador que aparece en el contexto genómico de la gentisaldehído deshidrogenasa no pertenece a la misma superfamilia del regulador transcripcional BetI. Los diferentes reguladores transcripcionales que forman parte del operón de estas enzimas, de nuevo sugieren que su inducción responde a compuestos diferentes y que por tanto estarían participando en rutas metabólicas diferentes y posiblemente tengan como sustratos a aldehídos también diferentes. Los resultados obtenidos de estos análisis sugieren fuertemente que la PA2125 no es una furfural deshidrogenasa, y que tampoco es una gentisaldehído deshidrogenasa o



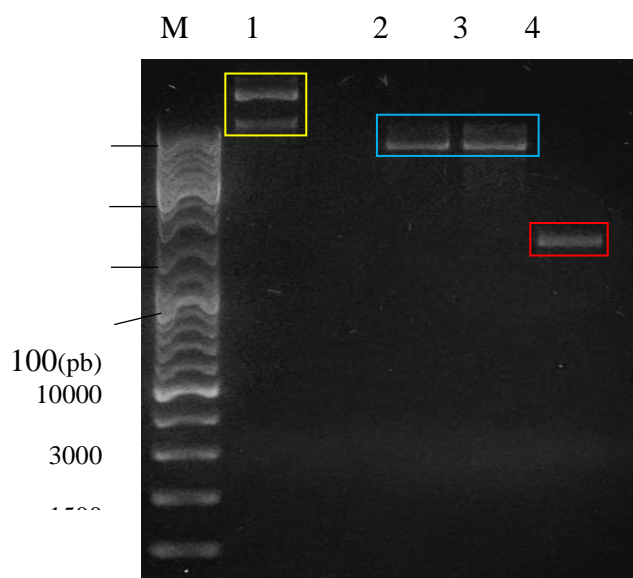
betaína aldehído deshidrogenasa. Pero no nos permiten descartar o afirmar que lo sea en forma concluyente o incluso que tenga otra actividad aún no reportada para las enzimas de su misma familia. Para investigar esto, necesitamos caracterizar bioquímicamente a la enzima PA2525 y estudiar experimentalmente su especificidad por sustrato para lo cual decidimos clonar el gen que la codifica a partir del DNA genómico de *P. aeruginosa* PAO1, insertarlo en un vector de expresión y obtener la enzima recombinante en células de *E. coli*.

## **VI.2 Clonación del gen PA2125**

El gen *pa2125* constituido por 1,449 pares de bases (pb) codifica un polipéptido de masa molecular de 51.3 kDa que corresponde a la secuencia de 482 residuos de amino ácidos de una subunidad de la enzima aldehído deshidrogenasa PA2125 de *P. aeruginosa*. El gen se clonó a partir del DNA genómico extraído de la bacteria y se amplificó mediante la técnica de PCR usando los oligonucleótidos diseñados con base en la secuencia del gen reportada en la base de datos NCBI. La amplificación del gen *pa2125* se confirmó por electroforesis en gel de agarosa (Figura 9), en el que se observaron bandas de aproximadamente 1,500 pb, que es el valor esperado para este gen.



**Figura 9.** Gel de agarosa en el que se muestra la amplificación por PCR del gen *pa2125*. Carriles: M, marcadores de masa molecular; 1, producto de PCR con un tamaño aproximado de 1500 pb (enmarcado en rojo).



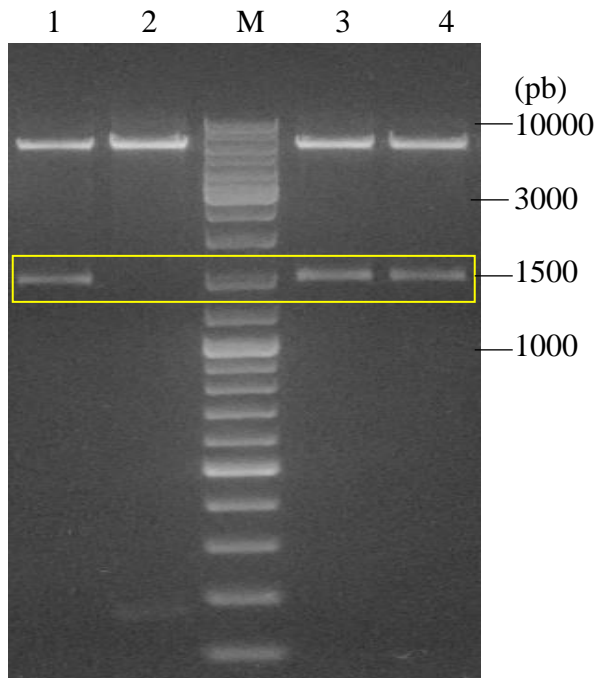
**Figura 10.** Gel de agarosa mostrando los productos de la digestión del vector *pET28b(+)* y del gen *pa2125* usando las enzimas *NdeI* y *HindIII*. Carriles: M, marcadores de masa molecular; 1 plásmido sin digerir a una concentración de 92 ng/μL (bandas enmarcadas en amarillo); 2 y 3, plásmido digerido a una concentración de 43 y 48 ng/μL respectivamente (bandas enmarcadas en azul); 4, amplicón de

*pa2125* digerido a una concentración de 38 ng / $\mu$ L (banda enmarcada en rojo). Se puede observar el diferente corrimiento del DNA del vector pET28b(+) sin digerir en dos tipos de conformación súper enrolladas (carril 1) de la conformación lineal del vector digerido (carriles 2 y 3).

El resultado de las digestiones del vector pET28b(+) antes de insertar el gen *pa2125* se muestra en el gel de agarosa de la Figura 10, carriles 2 y 3. Luego de comprobar que las digestiones fueron exitosas, se procedió a insertar el gen *pa2125* en el vector pET28b(+).

Con el plásmido construido, denominado pET28b-*pa2125*, se transformaron células competentes de *E. coli* XL10-Gold por medio de choque térmico y de las colonias que habían crecido en presencia de kanamicina se extrajo el plásmido al que, se le realizó una doble digestión con las enzimas de restricción *NdeI* y *HindIII*. El resultado de esta digestión demostró que el plásmido pET28-*pa2125* sí contenía un inserto del tamaño esperado, correspondiente al del gen *pa2125* (Figura 11).

A continuación, por medio de secuenciación y un alineamiento de secuencias por Blast, se comprobó que el inserto en el plásmido pET28b-*pa2125* era el gen *pa2125* reportado en la base de datos NCBI sin ningún cambio y que estaba en marco correcto de lectura (Anexo 4).



**Figura 11.** Gel de agarosa mostrando la digestión del plásmido pET28-pa2125 usando las enzimas *NdeI* y *HindIII*. Los plásmidos pET28-pa2125 fueron extraídos de colonias de células de *E. coli* XL10 Gold cultivadas en placas de agar LB con kanamicina (0.1 mg/mL). Carriles: M, marcadores de masa molecular; 1, 3 y 4, plásmido que contenía el inserto pa2125 (enmarcado en amarillo); 2 plásmido que no contenía el inserto pa2125.

### VI.3 Sobreexpresión del gen *pa2125* y producción de la enzima recombinante

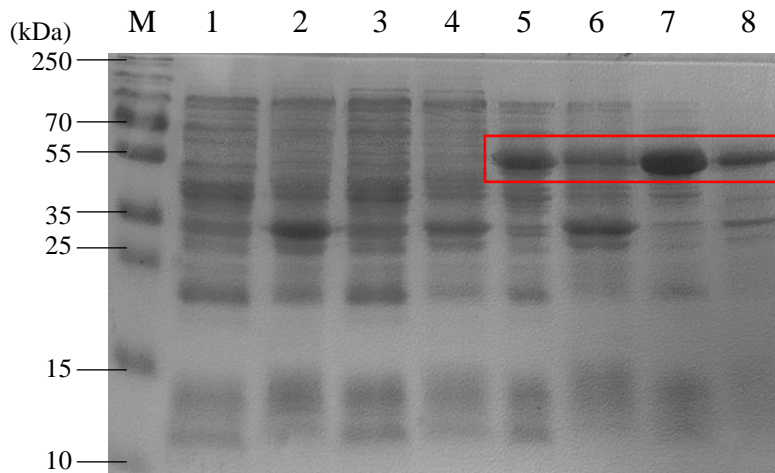
#### *PA2125* en células de *E. coli*

Se inocularon células de la cepa *E. coli* BL21 (DE3) Rosetta transformadas con el plásmido pET28b-pa2125 en medio LB estéril con kanamicina (0.1 mg/mL) y se indujeron con IPTG con el fin de realizar pruebas de expresión a pequeña escala (en 10 mL de medio de cultivo). Así mismo, como control se crecieron colonias de *E. coli* BL21 (DE3) Rosetta sin transformar. De todos los cultivos se tomaron alícuotas que se sometieron a SDS-PAGE

(Figura 5). Los marcadores de masa molecular permitieron identificar la banda de la enzima PA2125 en una posición que corresponde a una masa molecular aproximada de 55 kDa, un valor cercano al teórico de 51.3 kDa (bandas enmarcadas en rojo), que fue obtenido por medio de la herramienta *ProtParam* del portal de recursos bioinformáticos *Expasy* (<http://web.expasy.org/protparam/>).

Como se puede observar en la Figura 12 se obtuvo proteína PA2125 de forma soluble (carriles 5 y 7) e insoluble (carriles 6 y 8), aunque la proporción de la soluble fue mayor que la de la insoluble. Observamos que la proteína PA2125 se expresa aún en ausencia del inductor IPTG (carril 5), Dubendorff y Studier (1991) explican que cuando el represor lac (LacI) se une a una región de DNA del operador lac (contenido en el plásmido pET28b-pa2125) ubicada aproximadamente 15 pb hacia abajo del inicio de la secuencia del promotor T7 RNA inhibe fuertemente la transcripción por la T7 RNA polimerasa, mientras que si se une a una región centrada 21 pb hacia abajo del inicio del RNA del promotor T7 la inhibición se reduce sustancialmente. La baja represión sugiere que una vez iniciada la cadena, la unión del represor puede no ser muy efectiva para bloquear la elongación, y esto podría ser porque hay mutaciones en la región del operador del represor LacI.

A partir de las células por cada litro de cultivo, y a partir de ellas, tras la purificación se obtuvieron entre 50 y 60 mg de proteína PA2125 pura.

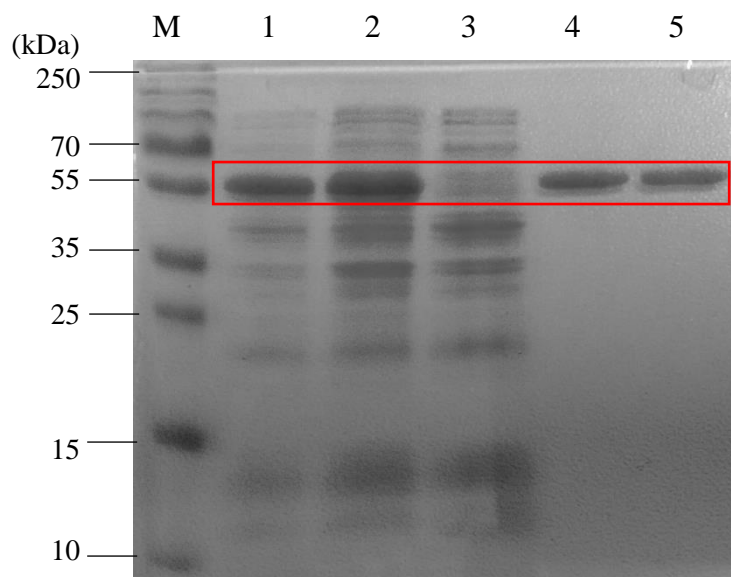


**Figura 12. Gel SDS-PAGE mostrando la producción de la enzima PA2125 recombinante en células de *E. coli* BL21 (DE3) Rosetta.** Carriles: M, marcadores de masa molecular; 1 y 2, sobrenadante y precipitado, respectivamente, del extracto crudo de las células sin transformar; 3 y 4 sobrenadante y precipitado, respectivamente, del extracto crudo de las células sin transformar y tratadas con IPTG; 5 y 6 sobrenadante y precipitado, respectivamente, del extracto crudo de las células transformadas y sin tratar con IPTG; 7 y 8 sobrenadante y precipitado, respectivamente, del extracto crudo de las células transformadas y tratadas con IPTG . Las bandas enmarcadas en rojo de aproximadamente 55 kDa corresponden a la enzima PA2125 producida. En cada carril se aplicaron 25 ug de proteína total.

#### **VI.4 Purificación a homogeneidad de la enzima PA2125**

La purificación se llevó a cabo en un solo paso que consistió en una cromatografía de afinidad a Ni. Se recogieron fracciones de 10 mL, de las cuales se reunieron las 10 primeras que eran las que presentaron actividad benzaldehído deshidrogenasa y tenían la mayor concentración de proteína. La pureza de la enzima PA2125 eluida se comprobó al observar una única banda de proteína en SDS-PAGE (Figura 13). Estas fracciones reunidas se

aplicaron a una columna de intercambio aniónico con el fin de eliminar de forma sencilla el imidazol de la preparación de enzima pura, evitando la agregación de ésta que ocurre cuando la enzima se concentra mediante lavados en tubos Centricón. Después de eluir, se reunieron las fracciones 2, 3 y 4, que fueron las únicas en presentar actividad benzaldehído deshidrogenasa, se hicieron alicuotas y se guardaron a -80 °C para su posterior uso. Los resultados de una purificación típica se muestran en la Figura 13 y en la Tabla 3.



**Figura 13. Gel SDS-PAGE mostrando el resultado de la purificación de la enzima PA2125.** Carriles: M, marcadores de masa molecular; 1, precipitado del extracto crudo (25 µg de proteína total aplicados); 2, sobrenadante del extracto crudo (25 µg de proteína total aplicados) 3, fracción de proteína soluble no unida a la columna de afinidad a Ni (25 µg de proteína total aplicados); 4, fracciones eluídas de la columna de afinidad a Ni y que presentaron mayor actividad enzimática y mayor concentración de proteína (3 µg de proteína total aplicados); 5, fracciones eluídas de la columna de intercambio iónico que presentaron mayor actividad enzimática y mayor concentración de proteína (3 µg de proteína total aplicados). Las bandas enmarcadas en rojo de aproximadamente 55 kDa corresponden a la enzima PA2125 producida.

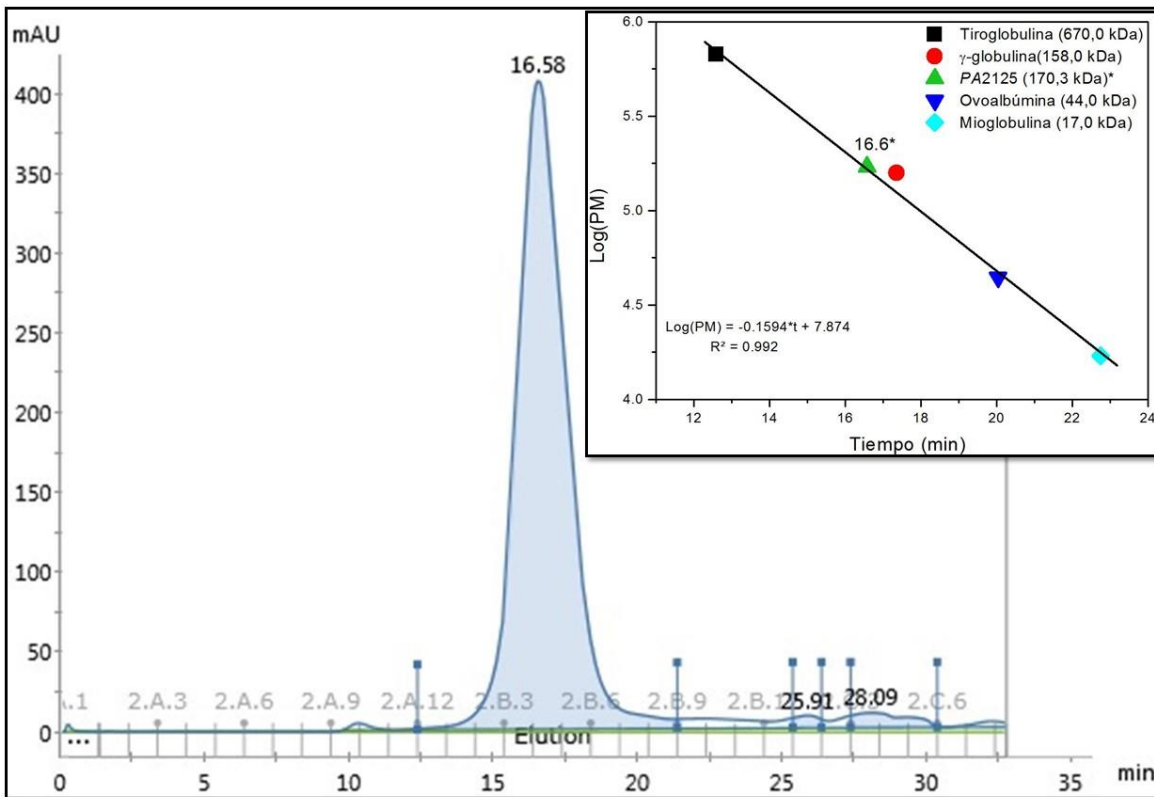
**Tabla 3. Resultados de la purificación de la enzima PA2125.** Se inició con 4 L del cultivo de bacterias en fase estacionaria. La actividad se midió con 2 mM de benzaldehído y 2 mM de NAD<sup>+</sup>.

Paso de purificación	Proteína total (mg)	Actividad total (Unidades)	Actividad específica (U/mg proteína)	Purificación (veces)	Rendimiento (%)
Extracto crudo	537	1,560	2.91	1	100
Cromatografía de afinidad a Ni	232	935	4.03	1.4	43

### VI.5 Estado oligomérico de la enzima PA2125

La cromatografía de exclusión molecular de la enzima recombinante PA2125 pura mostró un pico muy simétrico en un tiempo de retención de 12.41 min (Figura 11), lo que indica la pureza de la proteína que eluye en este pico. De acuerdo con la curva de calibración utilizada (inserto Figura 14) este pico corresponde a una proteína de 170.3 kDa, y teniendo en cuenta que la masa molecular teórica del monómero de la proteína PA2125 (51.3 kDa) parece corresponder a un trímero. Sin embargo, a la fecha no se ha reportado una estructura cristalográfica de una ALDH que presente una estructura cuaternaria trimérica, la que además no es consistente con el arreglo de dímeros o dímero de dímeros que presentan estas enzimas. Sí se conoce que algunas ALDHs tetraméricas presentan este patrón de elución anormal en SEC (Mori *et al*, 1980; Falkenberg y Strøm, 1990; Velasco-García *et al.*, 1999), por lo que asumimos que la estructura nativa de la enzima es tetramérica.



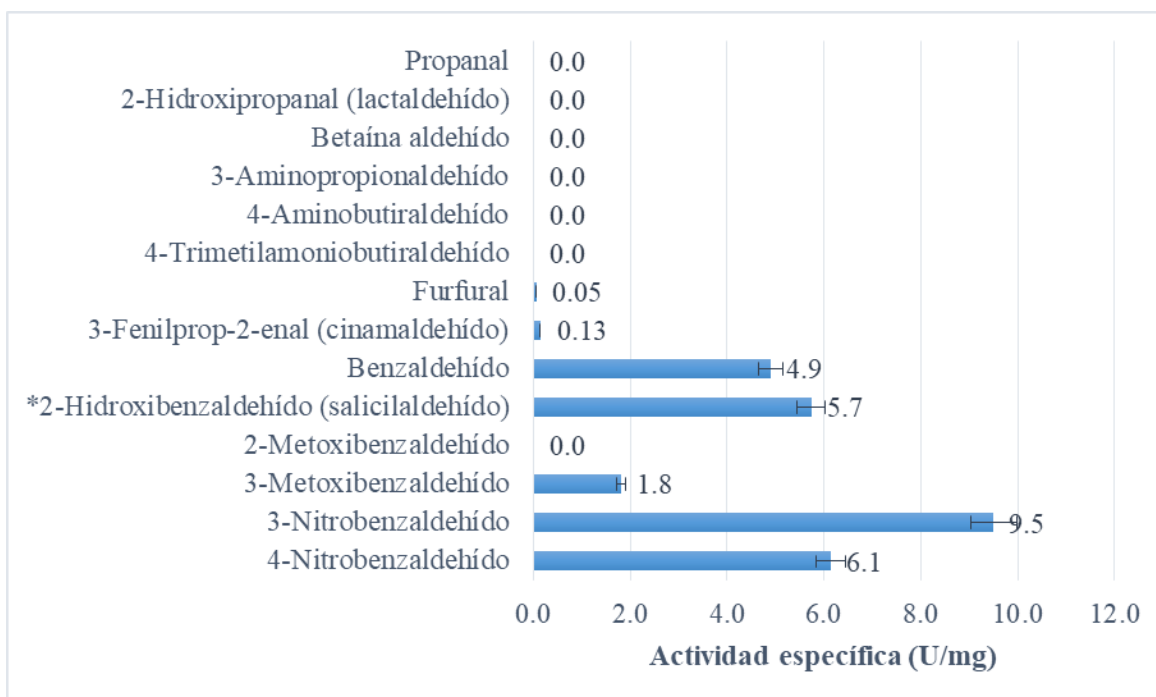


**Figura 14. Perfil de elución de la enzima PA2125 de la cromatografía de exclusión molecular.** La cantidad de proteína total aplicada fue 5 mg. La absorbancia a 280 nm está dada en miliunidades de absorbancia (mAU). Inserto: curva de calibración obtenida usando estándares de masa molecular.

### VI.6 Especificidad por el sustrato aldehído

Después de obtener pura la forma recombinante de la enzima PA2125, se estudió su especificidad para tratar de conocer cuál o cuáles aldehídos pudieran ser sus sustratos fisiológicos. Empezamos ensayando a la betaína aldehído (BAL) debido a que PA2125 fue anotada originalmente como una betaína aldehído deshidrogenasa (BADH) y a que en su contexto genómico hay una enzima alcohol deshidrogenasa como en el de la enzima PaBADH (PA5373), la cual está bien caracterizada como una betaína aldehído deshidrogenasa (Nagasawa *et al.*, 1976; Velasco-García *et al.*, 2006). Así mismo, se

ensayaron otros aminoaldehídos como el 4-N-trimetil-aminobutiraldehído (TMABAL), 4-aminobutiraldehído (ABAL) y 3-aminopropionaldehído (APAL), todos sustratos de otras enzimas ALDH que poseen actividad BADH (Díaz-Sánchez *et al.*, 2012). Se evaluó también la actividad enzimática con furfural debido al resultado obtenido con los alineamientos múltiples en los que encontramos una ALDH de *P. putida* (PsfA) que presenta un 47 % de identidad de secuencia de aminoácidos con PA2125, lo que nos indicaba que ambas enzimas pertenecen a la misma familia de enzimas de la superfamilia de ALDHs. Además, se hicieron ensayos de actividad enzimática con BZAL y compuestos derivados del BZAL porque, como se mencionó anteriormente, también encontramos en estos estudios una enzima que pertenece a la misma familia ALDH que PA2125 y que parece usar como sustrato al gentisaldehído, compuesto que no se usó porque no se contó con el reactivo. Para comparar los datos de velocidad obtenidos con todos estos sustratos y poder estimar su relevancia fisiológica, todos los ensayos enzimáticos se realizaron a concentraciones de los aldehídos 2 mM y de NAD<sup>+</sup> de 0.6 mM, que es la concentración intracelular de este cofactor estimada en de *E. coli* (Zhou *et al.*, 2011) y muy probablemente cercana a la que existe en *P. aeruginosa*. Los resultados obtenidos en estos ensayos de actividad se muestran en la Figura 15.

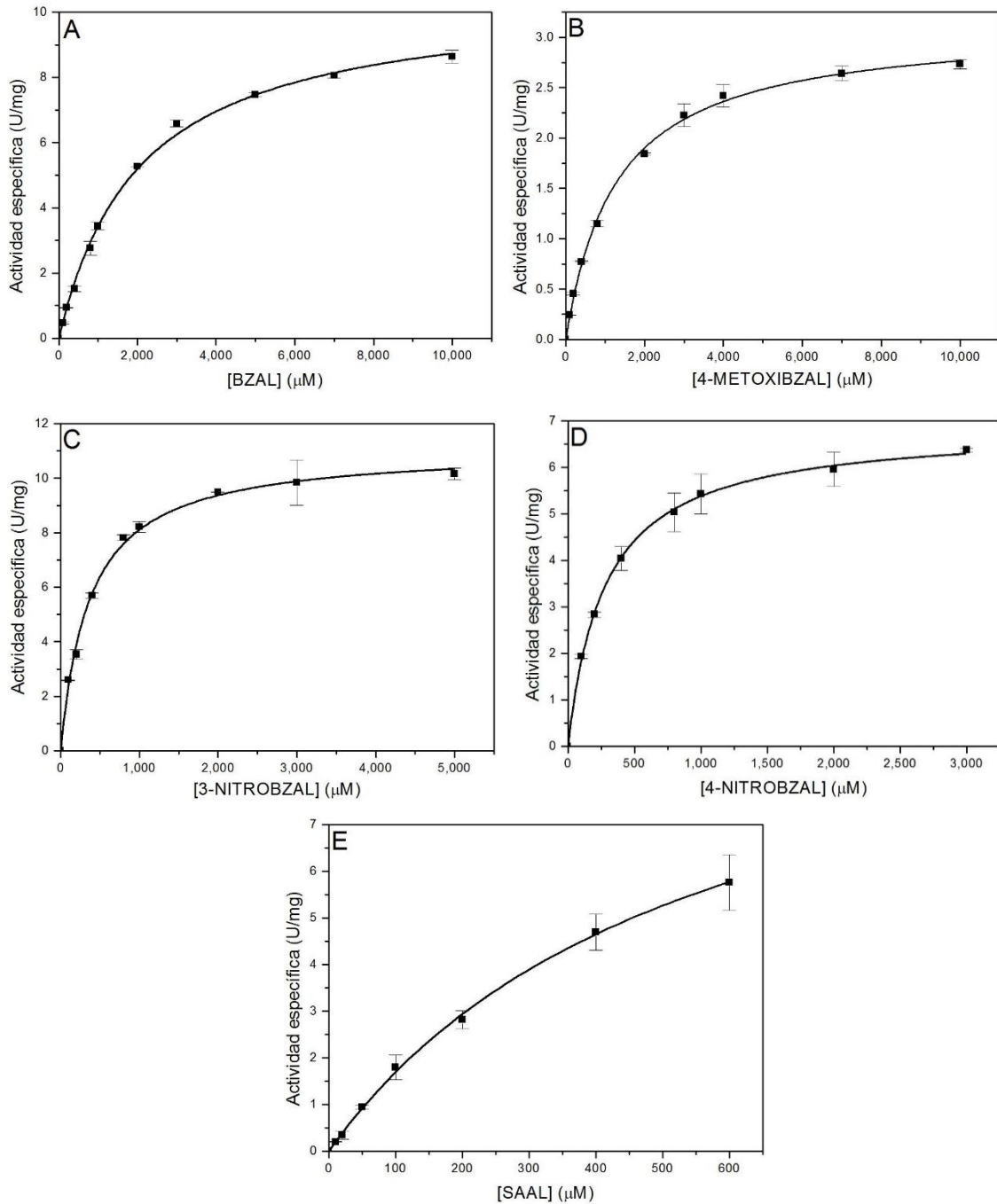


**Figura 15. Actividad de la enzima PA2125 con diversos aldehídos a concentración fisiológica de  $NAD^+$ .** Las velocidades iniciales se determinaron a una concentración de enzima de 20  $\mu\text{g/mL}$ , del aldehído 2 mM, y de  $NAD^+$  0.6 mM. Usando el ensayo descrito en la sección de Materiales y métodos. \*El ensayo con salicilaldehído se llevó a cabo a una concentración del aldehído de 600  $\mu\text{M}$ . Los valores son el promedio de los valores experimentales obtenidos y las barras verticales indican la desviación estándar del valor promedio de los datos de velocidad inicial de dos experimentos.

Como puede observarse, la enzima PA2125 no presentó actividad con aldehídos alifáticos ni con aminoaldehídos. Sí presentó actividad con furfural, 3-fenilprop-2-enal (cinamaldehído), aunque muy baja en estos dos casos, BZAL y compuestos sustituidos de BZAL. Bajo estas condiciones experimentales el mejor sustrato fue el 3-nitrobenzaldehído (3-NITROBZAL), seguido muy de cerca por el 4-nitrobenzaldehído (4-NITROBZAL) y el salicilaldehído (SAAL; 2-hidroxi-benzaldehído).

## VI.7 Parámetros cinéticos aparentes para benzaldehído y derivados

Para poder evaluar con mayor certeza cuál de todos los aldehídos que presentaron mayor actividad sería preferido por la enzima, se realizaron cinéticas de saturación usando a  $\text{NAD}^+$  como cofactor a una concentración fija de 0.6 mM. Cuando se realiza una curva de saturación variando la concentración de un sustrato pero manteniendo constante la concentración del otro sustrato, los parámetros cinéticos obtenidos  $apV$ ,  $apK_m$ ,  $apk_{cat}$  y  $apk_{cat} / K_m$  dependen de la concentración del sustrato cuya concentración permanezca constante, por lo que se les denomina parámetros cinéticos aparentes. Las curvas de saturación obtenidas se muestran en la Figura 16 y los parámetros cinéticos aparentes estimados a partir de los datos de velocidad inicial se incluyen en la Tabla 4. En base a las constantes de especificidad ( $apk_{cat} / K_m$ ) aparentes, proponemos cuál es el mejor sustrato aldehído a la concentración de  $\text{NAD}^+$  fisiológica.



**Figura 16. Cinéticas de saturación de la enzima PA2125 con los aldehídos como sustrato variable.** (A) Benzaldehído (BZAL), (B) 4-metoxibenzaldehído (4-METOXIBZAL), (C) 3-nitrobenzaldehído (3-NITROBZAL), (D) 4-nitrobenzaldehído (4-NITROBZAL) y (E) salicilaldehído (SAAL) como sustratos variables dejando la concentración de  $\text{NAD}^+$  constante a  $600 \mu\text{M}$ . En las gráficas, los puntos son el promedio de los valores experimentales obtenidos

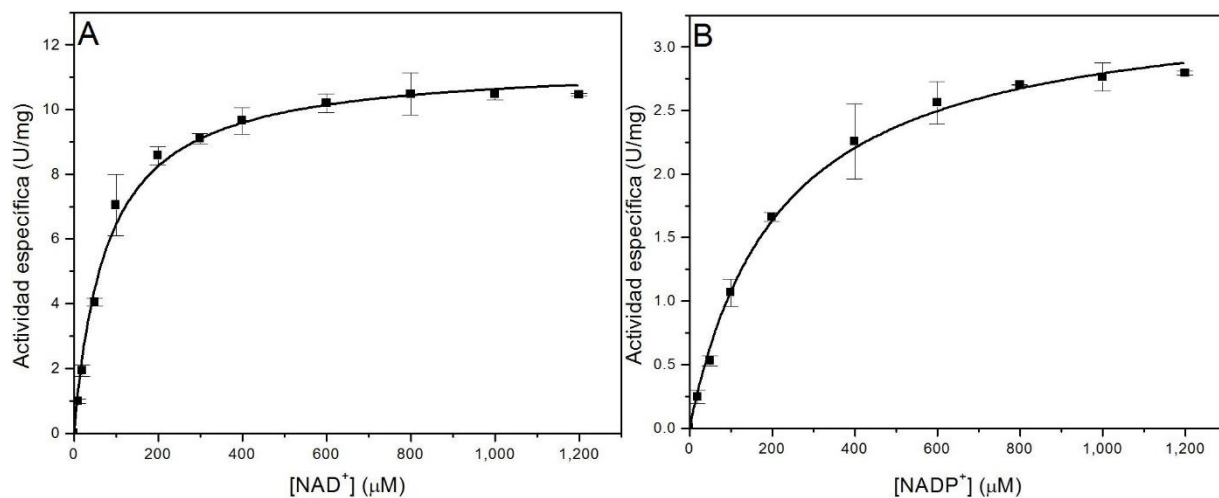
y las barras verticales indican la desviación estándar del valor promedio de los datos de velocidad inicial de dos experimentos. Las líneas trazadas son las calculadas a partir del mejor ajuste global de estos datos a la Ecuación 6 mediante regresión no lineal.

**Tabla 4. Parámetros cinéticos aparentes de la enzima PA2125 obtenidos con benzaldehído y aldehídos derivados.** La concentración de NAD<sup>+</sup> se mantuvo constante a 0.6 mM. Los valores de los parámetros cinéticos se obtuvieron por el ajuste de los datos experimentales mostrados en la Figura 16. El símbolo ± indica el error (desviación estándar) de los valores estimados.

Sustrato variable	$apK_m$ ( $\mu\text{M}$ )	$apV$ ( $\mu\text{mol}\cdot\text{min}^{-1}\cdot\text{mg}^{-1}$ )	$apk_{\text{cat}}$ ( $\text{s}^{-1}$ )	$apk_{\text{cat}}/K_m$ ( $1 \times 10^3 \text{ M}^{-1} \text{ s}^{-1}$ )
BZAL	2,047 ± 114	10.5 ± 0.2	9.0	4.4
4-METOXIBZAL	1,284 ± 63	3.1 ± 0.0	2.7	2.1
3-NITROBZAL	368 ± 23	11.1 ± 0.2	9.5	25.8
4-NITROBZAL	277 ± 11	6.9 ± 0.1	5.9	21.3
SAAL	559 ± 43	11.1 ± 0.5	9.8	17.5

De acuerdo con los valores aparentes de  $apk_{\text{cat}}/K_m$  dados en la Tabla 4, se confirma que a una concentración fisiológica del cofactor 3-NITROBZAL y 4-NITROBZAL son los mejores sustratos para PA2125, seguidos por SAAL con valores de  $apk_{\text{cat}}/K_m$  cercanos, y por último BZAL y 4-METOXIBZAL, este último con un valor de  $apk_{\text{cat}}/K_m$  un orden de magnitud por debajo del estimado para 3-NITROBZAL. Estos datos nos están indicando que la enzima prefiere sustratos aldehídos aromáticos sustituidos con grupos polares. Igualmente, para determinar cuál era el cofactor preferido por la enzima PA2125 se

hicieron cinéticas de saturación por  $\text{NAD}^+$  y  $\text{NADP}^+$  a una concentración fija de 10 mM de BZAL. Los resultados se muestran en la Figura 15 y en la Tabla 5.



**Figura 17. Cinéticas de saturación de la enzima PA2125 con  $\text{NAD(P)}^+$ .**  $\text{NAD}^+$  (A),  $\text{NADP}^+$  (B), como sustratos variables dejando la concentración de BZAL constante a 10 mM. En las gráficas, los puntos son el promedio de los valores de velocidad inicial obtenidos experimentalmente; las barras verticales indican la desviación estándar del valor promedio de estos datos experimentales. Las líneas trazadas son las calculadas a partir del mejor ajuste global de los datos a la Ecuación 6 mediante regresión no lineal.

**Tabla 5. Parámetros cinéticos aparentes de la enzima PA2125 obtenidos con  $\text{NAD}^+$  y  $\text{NADP}^+$ .** Los valores de los parámetros cinéticos se obtuvieron por el ajuste de los datos experimentales mostrados en Figura 15. El símbolo  $\pm$  indica el error (desviación estándar) de los valores estimados.

Sustrato variable	$apK_m$ ( $\mu\text{M}$ )	$apV$ ( $\mu\text{mol}\cdot\text{min}^{-1}\cdot\text{mg}^{-1}$ )	$apk_{\text{cat}}$ ( $\text{s}^{-1}$ )	$apk_{\text{cat}}/K_m$ ( $1 \times 10^5 \text{ M}^{-1} \text{ s}^{-1}$ )
$\text{NAD}^+$	$78 \pm 7$	$11.4 \pm 0.2$	9.8	1.2
$\text{NADP}^+$	$215 \pm 15$	$3.4 \pm 0.1$	2.8	0.13

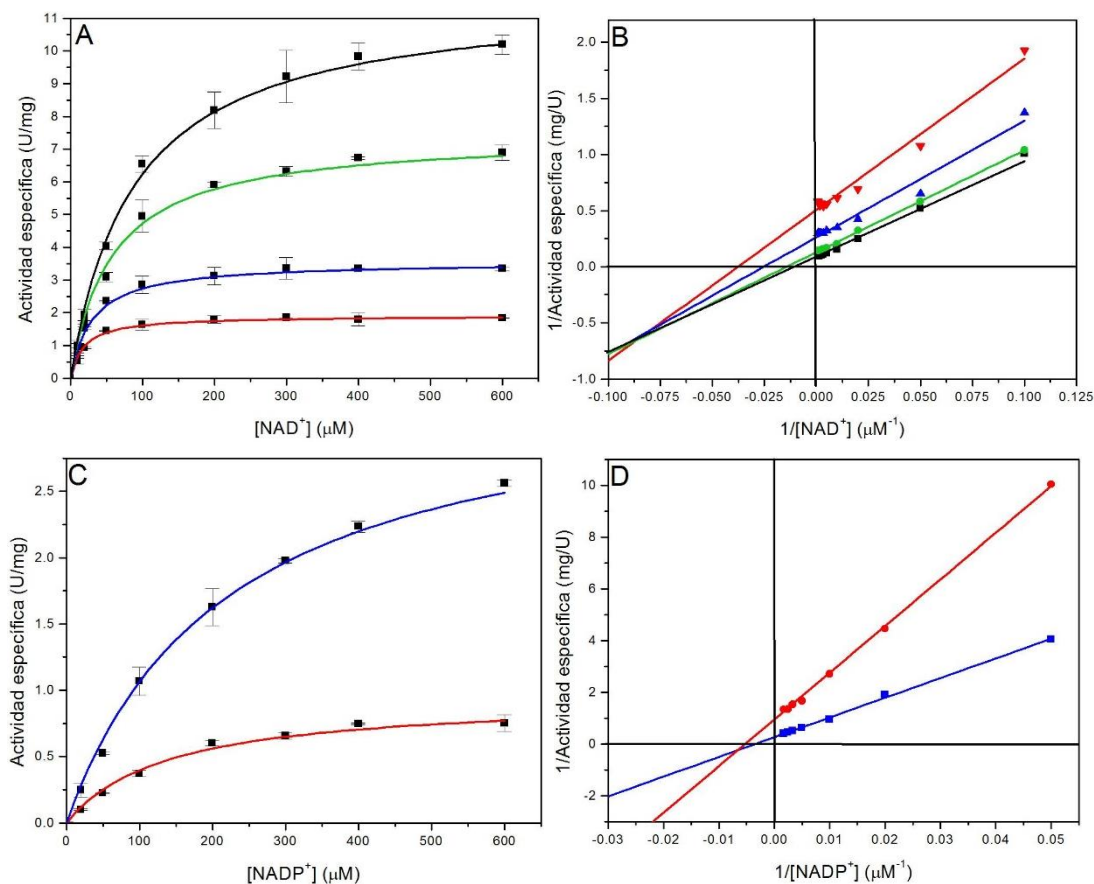
En cuanto a la especificidad de la enzima por el cofactor, los valores de  $ap_{k_{cat}}/K_m$  incluidos en la Tabla 5 indican que el  $NAD^+$  es mucho mejor sustrato que  $NADP^+$ , con un valor de  $ap_{k_{cat}}/K_m^{NAD^+}$  diez veces superior al valor de  $ap_{k_{cat}}/K_m^{NADP^+}$ . El mecanismo cinético de la enzima PA2125 se investigó determinando los patrones de velocidad inicial de las reacciones  $NAD^+$ - y  $NADP^+$ -dependientes con BZAL como sustrato.

### **VI.8 Mecanismo cinético de pa2125 usando benzaldehído como sustrato**

El mecanismo cinético de la enzima PA2125 se investigó determinando los patrones de velocidad inicial de las reacciones  $NAD^+$ - y  $NADP^+$ -dependientes con BZAL como sustrato (Figura 18). La concentración del aldehído se varió de 1 a 10 mM, y la concentración de  $NAD^+$  se varió en un intervalo de 0 a 1,200  $\mu M$ . Cuando los datos se grafican como dobles recíprocos con  $NAD^+$  como sustrato variable y BZAL como fijo variable se puede observar que las líneas se cruzan a la izquierda del eje de ordenadas y debajo del eje de abscisas, es decir en el tercer cuadrante, lo que indica que el mecanismo cinético es secuencial, como se ha descrito para la mayor parte de las ALDHs cuyo mecanismo cinético ha sido estudiado (Feldman y Weiner, 1972; Valenzuela-Soto y Muñoz-Clares, 1993; Figueroa-Soto y Valenzuela-Soto, 2000). Los datos obtenidos se ajustaron bien a la ecuación correspondiente a un mecanismo cinético Bi Bi ordenado en estado estacionario (Ecuación 7) en el que consideramos que  $NAD^+$  es el primer sustrato que se une a la enzima seguido por el aldehído, igualmente como se ha descrito para otras ALDHs (Feldman y Weiner, 1972; Valenzuela-Soto y Muñoz-Clares, 1993; Velasco-García *et al.*, 2000). Los parámetros cinéticos estimados se dan en la Tabla 6. En el patrón de velocidad inicial las líneas se cruzan en el tercer cuadrante debido a que la constante de disociación para  $NAD^+$



del complejo E-NAD<sup>+</sup> ( $K_{ia}$ ) es menor que la  $K_m^{NAD^+}$ . Aunque con sólo dos líneas, igual comportamiento se observa en el caso del NADP<sup>+</sup> (Figura 18), por lo que pensamos que el mecanismo cinético es igual con cualquiera de los dos cofactores.



**Figura 18. Patrones de velocidad inicial de la reacción de la enzima PA2125 con BZAL.**

A. Gráficas directas de la reacción NAD<sup>+</sup>-dependiente. B. Gráficas de dobles recíprocos correspondientes. La coenzima NAD<sup>+</sup> se varió en el intervalo de 10 a 600 μM a las concentraciones fijas variables de BZAL: 1 (▼); 2 (▲), 5 (●) y 10 mM (■); C. Gráficas directas de la reacción NADP<sup>+</sup>-dependiente. D. Gráficas de dobles recíprocos correspondientes. La coenzima NADP<sup>+</sup> se varió en el intervalo de 20 a 600 μM a las concentraciones fijas variables de BZAL: 1 (●) y 10 (■) mM. En las gráficas directas, los puntos son el promedio de los valores de velocidad inicial obtenidos experimentalmente; las barras verticales indican la desviación estándar del valor promedio de estos datos experimentales. Las líneas

trazadas son las calculadas a partir del mejor ajuste global de estos datos a la Ecuación 8 mediante regresión no lineal.

**Tabla 6. Parámetros cinéticos de la enzima PA2125 obtenidos con BZAL como sustrato.**

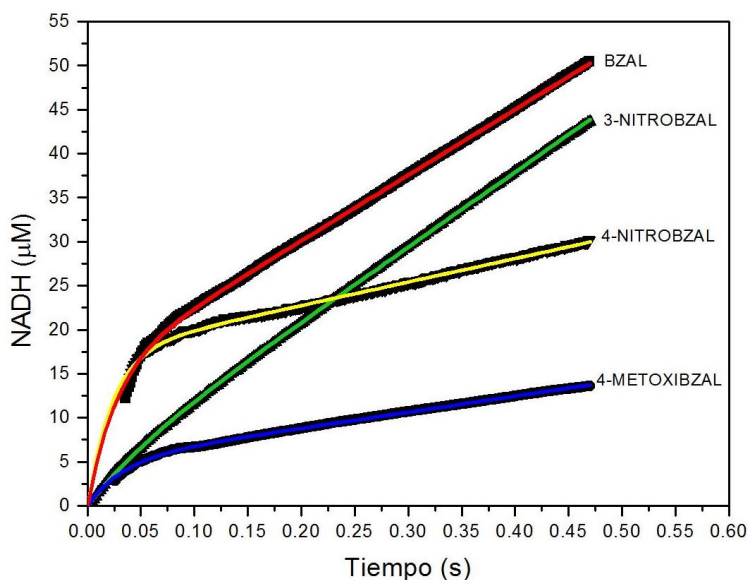
Los valores de los parámetros cinéticos se obtuvieron por el ajuste de los datos experimentales mostrados en la Figura 16. El símbolo  $\pm$  indica el error (desviación estándar) de los valores estimados.

Parámetro cinético	Valor estimado	
	NAD <sup>+</sup>	NADP <sup>+</sup>
$V$ ( $\mu\text{mol}\cdot\text{min}^{-1}\cdot\text{mg}^{-1}$ )	$27 \pm 1$	$4.8 \pm 0.2$
$k_{\text{cat}}$ ( $\text{s}^{-1}$ )	23.4	4.1
$K_{\text{ia}}$ ( $\mu\text{M}$ )	$5.0 \pm 3.0$	$107 \pm 29$
$K_m^{\text{NAD(P)}^+}$ ( $\mu\text{M}$ )	$198 \pm 14$	$265 \pm 23$
$K_m^{\text{BZAL}}$ (mM)	$13.4 \pm 0.7$	$4.0 \pm 0.3$
$k_{\text{cat}} / K_m^{\text{NAD(P)}^+}$ ( $\text{M}^{-1} \cdot \text{s}^{-1}$ )	$1.18 \times 10^5$	$1.54 \times 10^4$
$k_{\text{cat}} / K_m^{\text{BZAL}}$ ( $\text{M}^{-1} \cdot \text{s}^{-1}$ )	$1.75 \times 10^3$	$1.01 \times 10^3$

## VI.9 Cinética rápida

Para determinar cuál es el paso limitante de la reacción catalizada por la enzima PA2125, ya sea de la etapa de acilación o de desacilación, y para entender así mejor los resultados de cinética de estado estacionario sobre la especificidad por sustrato de esta enzima, se llevaron a cabo ensayos de cinética rápida en un equipo de flujo detenido. Se usaron aquellos aldehídos con los que PA2125 presentó mayor actividad: BZAL (10 mM), 4-

METOXIBZAL (10 mM), 3-NITROBZAL (5 mM) y 4-NITROBZAL (3 mM), todos ellos a concentraciones saturantes, y como cofactor  $\text{NAD}^+$  a la concentración de 0.6 mM; la concentración de sitios activos de la enzima *PA2125* fue de 15.8  $\mu\text{M}$ . Los cursos temporales obtenidos se muestran en la Figura 19, y los parámetros cinéticos estimados por medio de un ajuste de los datos a la Ecuación 8 se incluyen en la Tabla 7.



**Figura 19.** Cursos temporales de la actividad de la enzima *PA2125* en ensayos de cinética rápida. BZAL 10 mM (línea roja), 3-NITROBZAL 5 mM (línea verde), 4-NITROBZAL 3 mM (línea amarilla) y 4-METOXIBZAL 10 mM (línea azul). En la gráfica, los puntos son el promedio de los valores experimentales obtenidos en 5 ensayos; las líneas trazadas en color son las calculadas a partir del mejor ajuste global de estos datos a la Ecuación 8 mediante regresión no lineal.

Se observó un estallido (“burst”) en la producción de NADH durante el primer ciclo catalítico con los cuatro aldehídos ensayados (Figura 19). Los aldehídos BZAL, y 4-NITROBZAL presentan valores de amplitud de este estallido ( $A_{burst}$ ) similares entre sí, y muy cercanos al valor de concentración de sitios activos de *PA2125*, lo que nos indica que para estos sustratos la desacilación es mucho más lenta que la acilación.

**Tabla 7. Parámetros cinéticos del estado pre-estacionario de la enzima PA2125.** Los valores de los parámetros cinéticos se obtuvieron por el ajuste de los datos experimentales mostrados en la Figura 19. Las concentraciones fueron saturantes: BZAL (10 mM); 4-METOXIBZAL (10 mM); 3-NITROBZAL (5 mM); 4-NITROBZAL (3 mM); NAD<sup>+</sup> (0.6 mM). El símbolo ± indica el error (desviación estándar) de los valores estimados.

Sustrato	$A_{burst}$ ( $\mu\text{M}$ )	$k_{burst}$ ( $\text{s}^{-1}$ )	$v_{ss}$ ( $\mu\text{mol}\cdot\text{min}^{-1}\cdot\text{mg}^{-1}$ )	$apk_{cat}$ ( $\text{s}^{-1}$ )
BZAL	15.18 ± 0.03	39.1 ± 0.3	5.6	4.8
4-METOXIBZAL	5.15 ± 0.01	30.4 ± 0.2	1.4	1.2
3-NITROBZAL	3.90 ± 0.02	18.5 ± 0.2	6.4	5.5
4-NITROBZAL	17.40 ± 0.02	43.6 ± 0.3	2.0	1.7

Los valores de las constantes de velocidad  $k_{burst}$  para estos dos aldehídos son mayores que los valores de  $k_{burst}$  para 4-METOXIBZAL y para 3-NITROBZAL (Tabla 7) indicando que la acilación es más rápida con estos dos sustratos. También se observa que un grupo nitro (-NO<sub>2</sub>) o metoxi (-OCH<sub>3</sub>) en la posición 4 del anillo aromático del BZAL disminuye la velocidad del estado estacionario, que corresponde a la etapa de desacilación, cuando se la compara con la obtenida con BZAL o con 3-NITROBZAL. Estos resultados se discutirán en función de la naturaleza química del grupo sustituyente y de su posible posición en el sitio activo de la enzima sugerida en los experimentos de acoplamiento molecular cuyos resultados se describen más adelante.

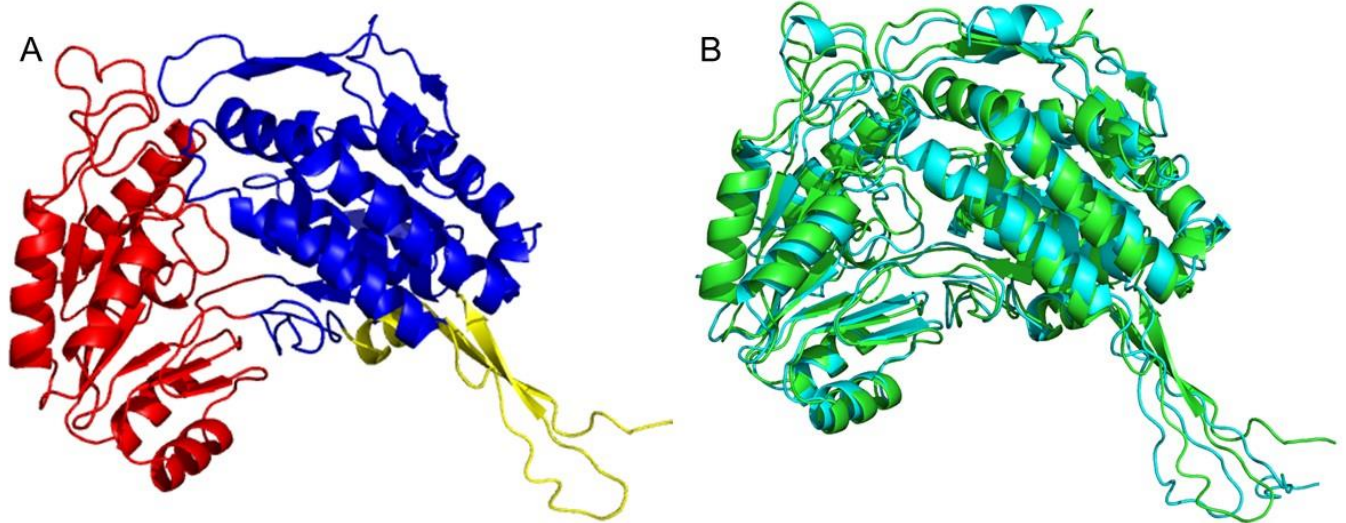
## VI.10 Modelo por homología de la estructura tridimensional de la enzima

### PA2125

Para modelar por homología la estructura tridimensional de la proteína PA2125 se buscó la estructura cristalográfica más adecuada para usarse como molde mediante los programas *Turingen's suit* (Söding *et al.*, 2005), HHpred (Söding, 2005) y HHsensor (Söding *et al.*, 2006), como se describió en la sección de Métodos. Con ambos programas obtuvimos prácticamente los mismos resultados de posibles moldes a usar. Teniendo en cuenta la integridad y el INDEX (Anexo 5) podrían usarse como moldes las estructuras de las enzimas: betaína aldehído deshidrogenasa de *Staphylococcus aureus* (SaBADH; CDD NCBI: cd07119)) de la que se seleccionaron las estructuras con códigos PDB 4QTO, 4MPB, 4Q92, 4ZWL, una aldehído deshidrogenasa de *Zea mays* (de la que se han reportado dos estructuras con códigos PDBS RF2C, y RF2F; CDD NCBI: cl11961), una aldehído deshidrogenasa de *Burkholderia cenopacia* (PDB 4YWE; CDD NCBI: cd07109) y la aldehído deshidrogenasa de *Brucella melitensis biovar abortus* 2308 (código PDB 3EK1; CDD NCBI: cd07103). El método Rd.HMM (Humphrey *et al.*, 1996) se utilizó para medir la habilidad (*fitness*) de estos posibles moldes para recuperar la secuencia de aminoácidos de PA2125 encontrándose que, con la excepción del PDB 3EK1, todos ellos podían hacerlo con significación estadística (bajo *E-value*) y puntuación HMM significativa positiva y cercana al 30% de la longitud de la secuencia proteica. De acuerdo con las puntuaciones Rd.HMM, ninguno de los moldes era perfecto para la secuencia de aminoácidos de PA2125, como se esperaba, pero el de mayor probabilidad de producir buenos modelos fue el PDB 4QTO, que por ello fue con el que se siguieron los procedimientos de modelado y dinámica molecular descritos en Métodos para obtener el modelo de la PA2125. Después de aproximadamente 10 ns de simulación de dinámica

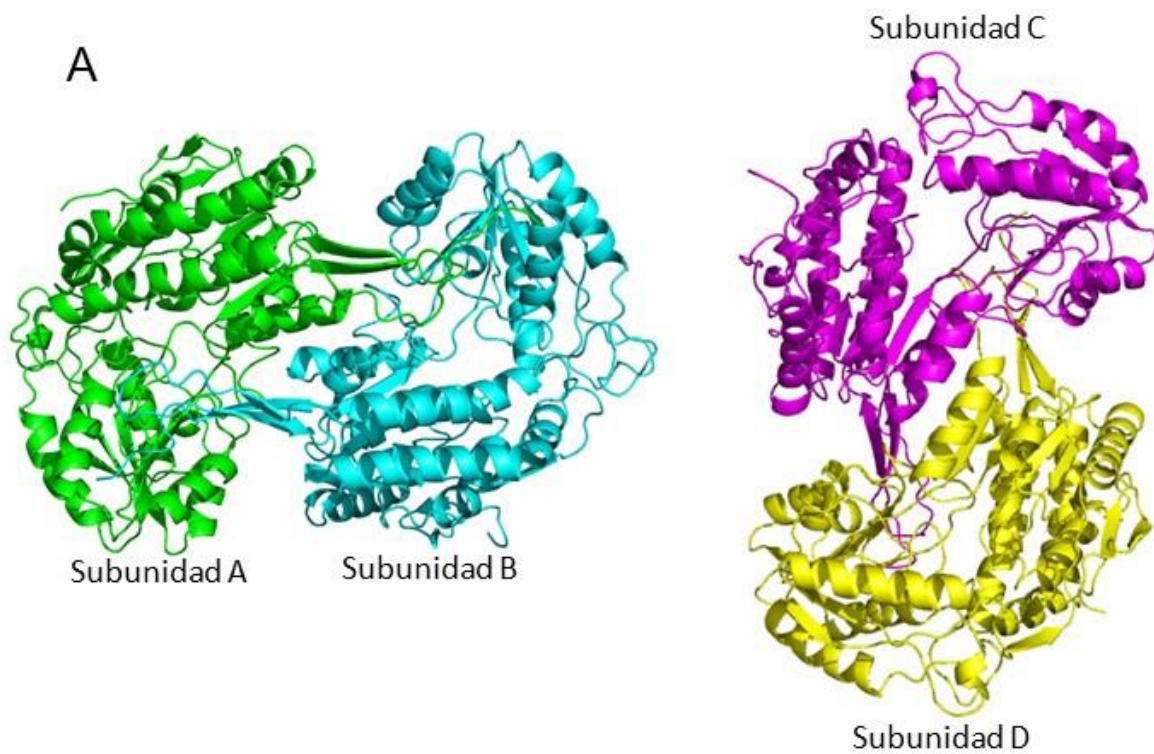
molecular, el RMSD y el radio de giro del modelo alcanzaron una meseta y mostraron principalmente fluctuaciones estocásticas. Los últimos 5 ns de la trayectoria se agruparon usando usando *clustering* jerárquico para producir 6 grupos de cada uno de los cuales se extrajo un representante cuya geometría se optimizó y se calificó con Rd.HMM.

Dos de los representantes de los 6 grupos fueron capaces de recuperar la secuencia de *PA2125* de la base de datos de secuencias no redundantes (NR) del NCBI como la primera primera opción, y el puntaje HMM (315) indicaba que se recuperaba el 65% de la secuencia de aminoácidos de la proteína *PA2125*, que es de 481 aminoácidos (Anexo 6). Por consiguiente, la estructura tridimensional predicha en el modelo por homología obtenido para *PA2125* y mostrado en las Figuras 20 y 21, aunque no es tan precisa como la estructura tridimensional que se obtiene por cristalografía experimentalmente, es altamente confiable, es decir, tiene una alta probabilidad de representar la estructura nativa de esta proteína. La mejora adicional de la predicción está limitada actualmente por el sesgo en el enfoque de campo de fuerza mecánica clásica y el coste computacional de simulaciones MD más largas para un sistema del tamaño del tetrámero de *PA2125*.

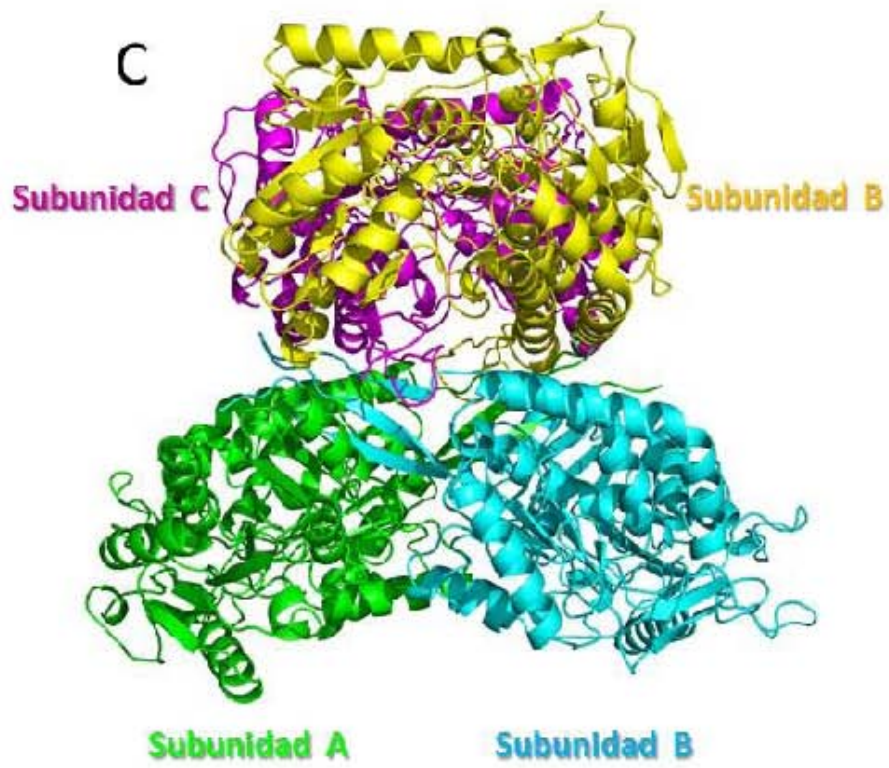
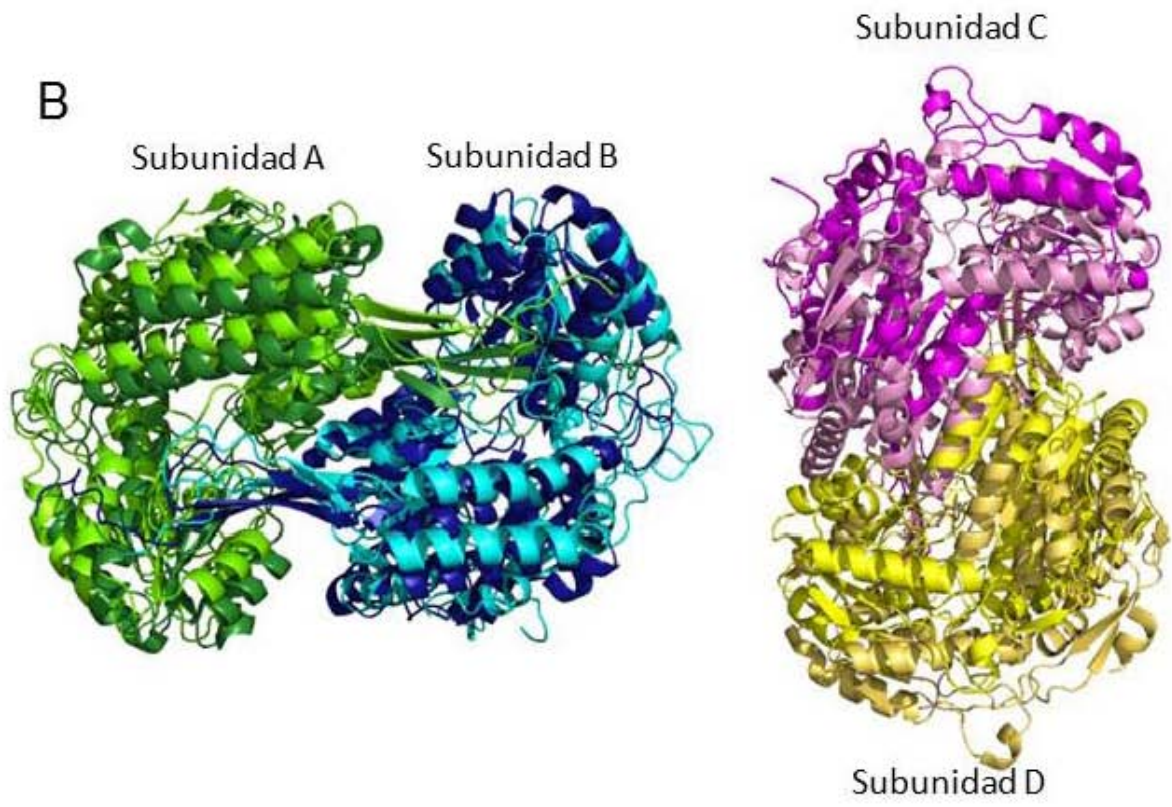


**Figura 20. Monómero del modelo por homología obtenido para la enzima PA2125. (A).**

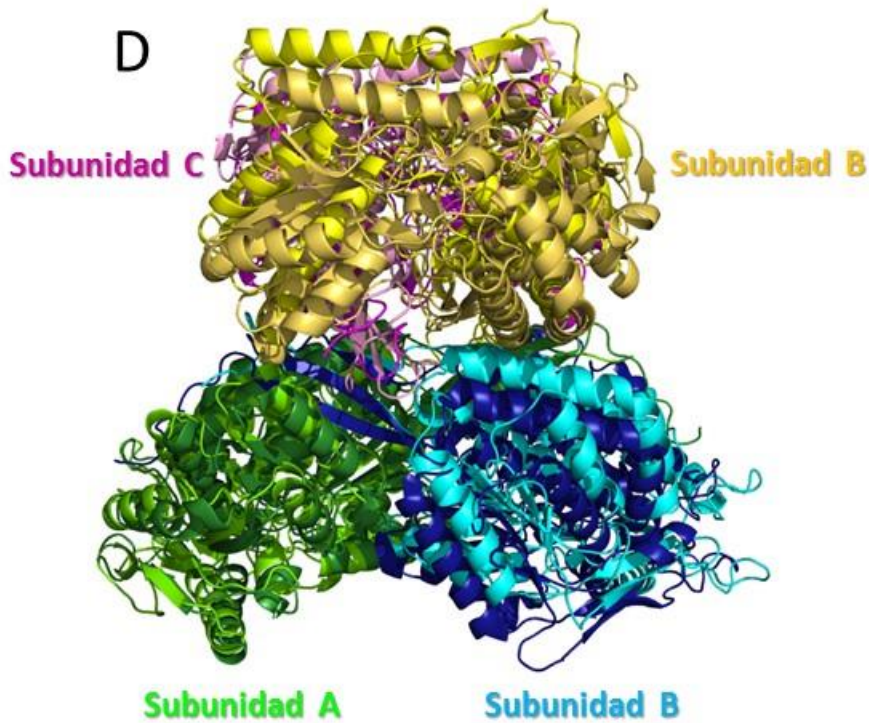
Representación del plegamiento y elementos de la estructura secundaria del monómero mostrando como lazos el dominio de catalítico (rojo), el dominio de unión al cofactor (azul) y el dominio de oligomerización (amarillo). (B). Comparación de la subunidad B del modelo por homología (color verde) con la subunidad B de la estructura cristalográfica usada como molde (*SaBADH*; código PDB 4QTO (color cian). Imágenes creadas con Pymol (<https://pymol.org/2/>).







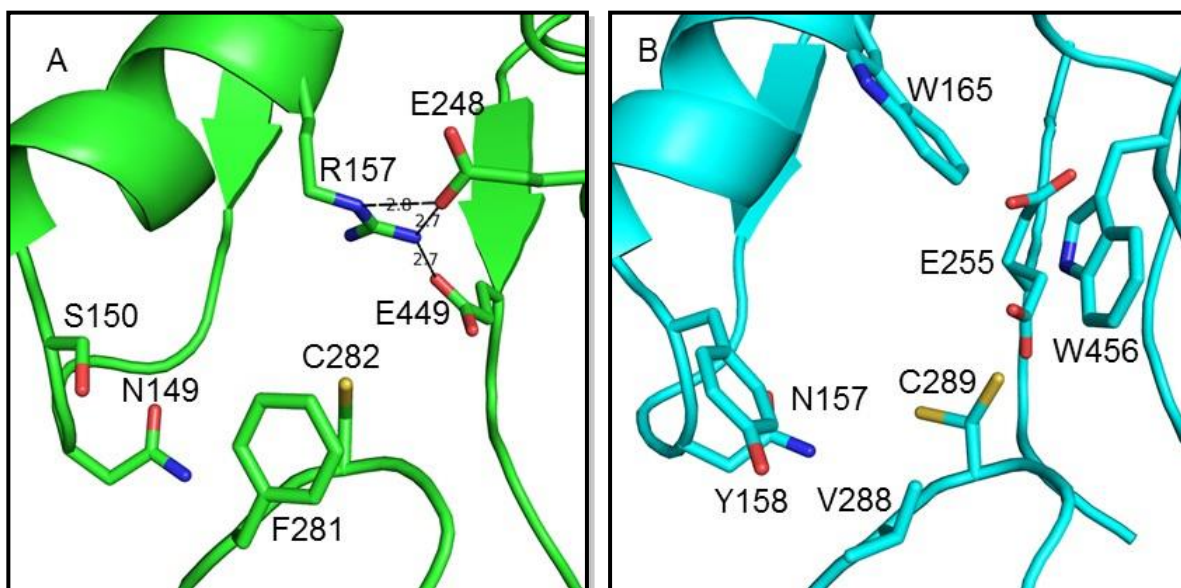




**Figura 21. Dímero y tetramero del modelo por homología obtenido para PA2125.** (A). Representación del plegamiento y elementos de la estructura secundaria de los dos dímeros del tetramero mostrando como lazos la subunidad A (verde), la subunidad B (cian), la subunidad C (violeta), y la subunidad D (amarillo); el dímero formado por subunidades C y D se giró 90° con respecto al dímero formado por subunidades A y B. (B). Dímeros del modelo por homología sobrepuesto con los dímeros de la enzima SaBADH con código PDB 4QTO. (C) Tetramero de modelo por homología de PA2125. (D) Superposición del tetramero de PA2125 con el tetramero del molde PDB 4QTO. Imágenes creadas en Pymol (<https://pymol.org/2/>).

La estructura del modelo por homología de la enzima PA2125 mostró a la cisteína catalítica (Cys282) en la conformación llamada de ataque, al Glu catalítico (Glu248) en la conformación “adentro” (González-Segura *et al.*, 2009) y a la asparagina catalítica (Asn149) en la posición encontrada hasta ahora en todas las estructuras cristalográficas, que

es la que le permite formar la cavidad del oxianión junto con el nitrógeno amídico de la Cys282. Estos tres residuos están conservados en todas las enzimas que pertenecen a las familias de las ALDHs hidrolíticas, incluida la de la PA2125 con cd07120 (Figura 7. Con este modelo por homología se confirmó la peculiaridad de las enzimas de esta familia de no tener residuos aromáticos en posiciones del sitio activo donde las enzimas de otras muchas familias los poseen, lo que habíamos encontrado por alineamiento de secuencias de aminoácidos y descrito anteriormente (Figura 7). En el caso de la PA2125, estos residuos son la Ser150, que está en lugar de una tirosina o fenilalanina existente en la mayoría de las ALDHs, la Arg157, que ocupa la posición equivalente a un residuo de triptófano en muchas de las ALDHs, y el Glu449, que corresponde a una fenilalanina en la mayoría de las ALDHs conocidas, aunque en algunas otras hay un triptófano o una histidina (Riveros-Rosas *et al.*, 2013). En la posición 281 de PA2125 hay un residuo de fenilalanina, que está conservado en la mayoría de las enzimas que forman la familia cd07120. Este residuo podría participar en determinar la especificidad por aldehídos aromáticos, como los acoplamientos moleculares indicaron. La disposición de todos estos residuos en el sitio activo se muestra en la Figura 22, comparando con los residuos aromáticos que están en las posiciones equivalentes en la estructura cristalográfica usada como molde.



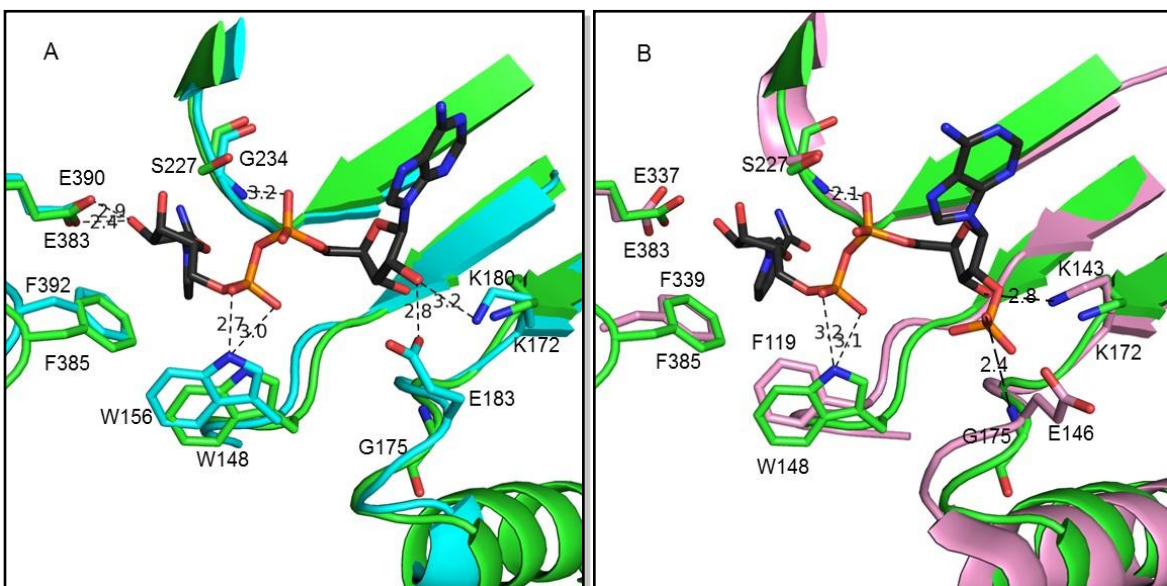
**Figura 22. Sitio activo del modelo por homología obtenido para la enzima PA2125. (A).**

Se muestran las cadenas laterales de los residuos del sitio activo de la subunidad B del modelo por homología de PA2125. (B) Cadenas laterales de los residuos equivalentes del sitio activo de la subunidad B de la estructura cristalográfica 4QTO de la enzima SaBADH usada como molde (cian). En el modelo de PA2125 la Cys282 catalítica está en la conformación “de ataque”, y el Glu248 catalítico en la conformación “adentro”; en la estructura 4QTO, la Cys289 catalítica está en las dos conformaciones “ataque” y “descanso” (ambas conformaciones formando un disulfuro mixto con el  $\beta$ -mercaptoetanol, el que no se muestra para simplificar la imagen), y el Glu255 catalítico en las conformaciones “adentro” e “intermedia”. En esta figura se muestra además la posición de los residuos que conforman la caja de aromáticos presente en la estructura molde (Tyr158, Trp165 y Trp456) y que en PA2125 son Ser150, Arg157 y Glu449. En el panel (A) se muestran en líneas discontinuas los puentes de hidrógeno entre los residuos Arg157 y Glu248, y entre Arg157 y Glu449. Los residuos de aminoácidos se muestran como varillas coloreadas por átomos (carbonos en verde en el modelo de PA2125 y en cian en el molde 4QTO, nitrógenos en azul, oxígenos en rojo, azufres en amarillo). Imagen creada con Pymol (<https://pymol.org/2/>).

En las enzimas con actividad BADH, o aminoaldehído deshidrogenasas, los residuos aromáticos forman una “caja de aromáticos” que participa en la unión del sustrato aldehído en el sitio activo mediante interacciones  $\pi$ -catión con el grupo de amonio cuaternario o amino. En el caso de otras enzimas ALDHs que oxidan a aldehídos aromáticos, como la benzaldehído deshidrogenasa de *P. putida* ATCC 12633 (CDD NCBI: cd07087), que es la única benzaldehído deshidrogenasa cristalizada a la fecha, estos residuos corresponden a Tyr143, Thr149 y Phe453 (PDB 5UCD; Zahniser *et al.*, 2017), de los cuales el residuo de tirosina y el de fenilalanina unen al anillo aromático del benzaldehído por medio de interacciones  $\pi$ -stacking. El tipo de interacciones que puedan hacer la Ser150, R157 y E449 con los mejores sustratos aldehídos que hemos encontrado en este estudio, lo investigamos por medio de experimentos de acoplamiento molecular (“docking”) que se describen más adelante.

En cuanto al sitio de unión a la coenzima (Figura 23), el modelo por homología de PA2125 muestra que el residuo de glicina (Gly175), que en los alineamientos se había mostrado que era el equivalente al residuo glutámico (Glu195) en *HsALDH2*, efectivamente ocupa la posición equivalente a este residuo (Glu183 en la estructura molde de la *SaBADH*), y, que por ser una glicina no puede interaccionar con la ribosa de la adenosina, como sí lo hace el residuo de glutámico en esta posición en otras enzimas. Se sabe que ésta es la posición que en gran parte determina la especificidad por  $\text{NAD}^+$  o  $\text{NADP}^+$  (González-Segura *et al.*, 2013). En el caso de la PA2125, la presencia de un residuo pequeño sin carga es consistente con el uso tanto de  $\text{NAD}^+$  como de  $\text{NADP}^+$ , lo que concuerda con los datos experimentales descritos anteriormente, aunque la enzima prefiere al  $\text{NAD}^+$  por razones no aparentes a simple vista.

Para inferir las posibles interacciones del sitio de unión al cofactor con el NAD(P)<sup>+</sup> superpusimos el modelo por homología de la enzima PA2125 con la estructura cristalográfica de la enzima SaBADH en complejo con NAD<sup>+</sup> (código PDB 4NEA), así como con la estructura cristalográfica de la enzima benzaldehído deshidrogenasa de *P. putida* ATCC 12633 en complejo con NADP<sup>+</sup> (código PDB 5UCD) (Figura 23). Los residuos que hacen interacciones importantes con el cofactor en la SaBADH (Trp156, Lys172 y Glu390) están conservados en la PA2125 (Trp148, Lys180, y Glu383, respectivamente) y estos podrían hacer las mismas interacciones que están haciendo en la enzima SaBADH. La diferencia importante es que la PA2125 tiene la Gly175 en lugar del Glu183 de la SaBADH, y por tanto no puede hacer el puente de hidrógeno que el grupo carboxilo de la cadena lateral de este residuo hace con el 2'-OH de la ribosa de la adenosina del NAD<sup>+</sup>. Esto puede explicar la relativamente baja afinidad de la PA2125 por este cofactor. En el caso de la unión del NADP<sup>+</sup>, la benzaldehído deshidrogenasa posee igualmente los residuos clave para unir al cofactor (Phe119, Lys143 y Glu337) y en la posición del Glu179 (numeración de *HsALDH2*) posee Glu146. El que la PA2125 una al NADP<sup>+</sup> se debe a que no existen impedimentos estéricos con el residuo de glicina (Gly175) y el grupo 2'-fosfato de la ribosa de la adenosina del NADP<sup>+</sup>. Por tanto, según el modelo por homología de la PA2125, tanto el NAD<sup>+</sup> como el NADP<sup>+</sup> pueden unirse bien al sitio de unión al cofactor; sin embargo este modelo no logra explicar por qué la enzima PA2125 presenta mayor afinidad por NAD<sup>+</sup> que por NADP<sup>+</sup>.



**Figura 23. Sitio de unión de la coenzima del modelo por homología de la enzima PA2125.** (A) Superposición del modelo por homología de PA2125 (átomos de carbono coloreados en verde) con la estructura cristalográfica de la enzima SaBADH (átomos de carbono coloreados cian) con NAD<sup>+</sup> unido (código PDB 4NEA). (B) Superposición del modelo por homología de PA2125 (átomos de carbono coloreados en verde) con la estructura cristalográfica de la enzima benzaldehído deshidrogenasa de *P. putida* (átomos de carbono coloreados en rosa) con NADP<sup>+</sup> unido (código PDB 5UCD). Los puentes de hidrogeno se muestran como líneas discontinuas. Las cadenas laterales de los aminoácidos están dibujadas como varillas coloreadas por tipo de átomo (oxígenos en rojo, nitrógenos en azul) y las cadenas principales como lazos. Los carbonos de los cofactores se muestran en color negro, fósforos en naranja, oxígenos en rojo y nitrógenos en azul.

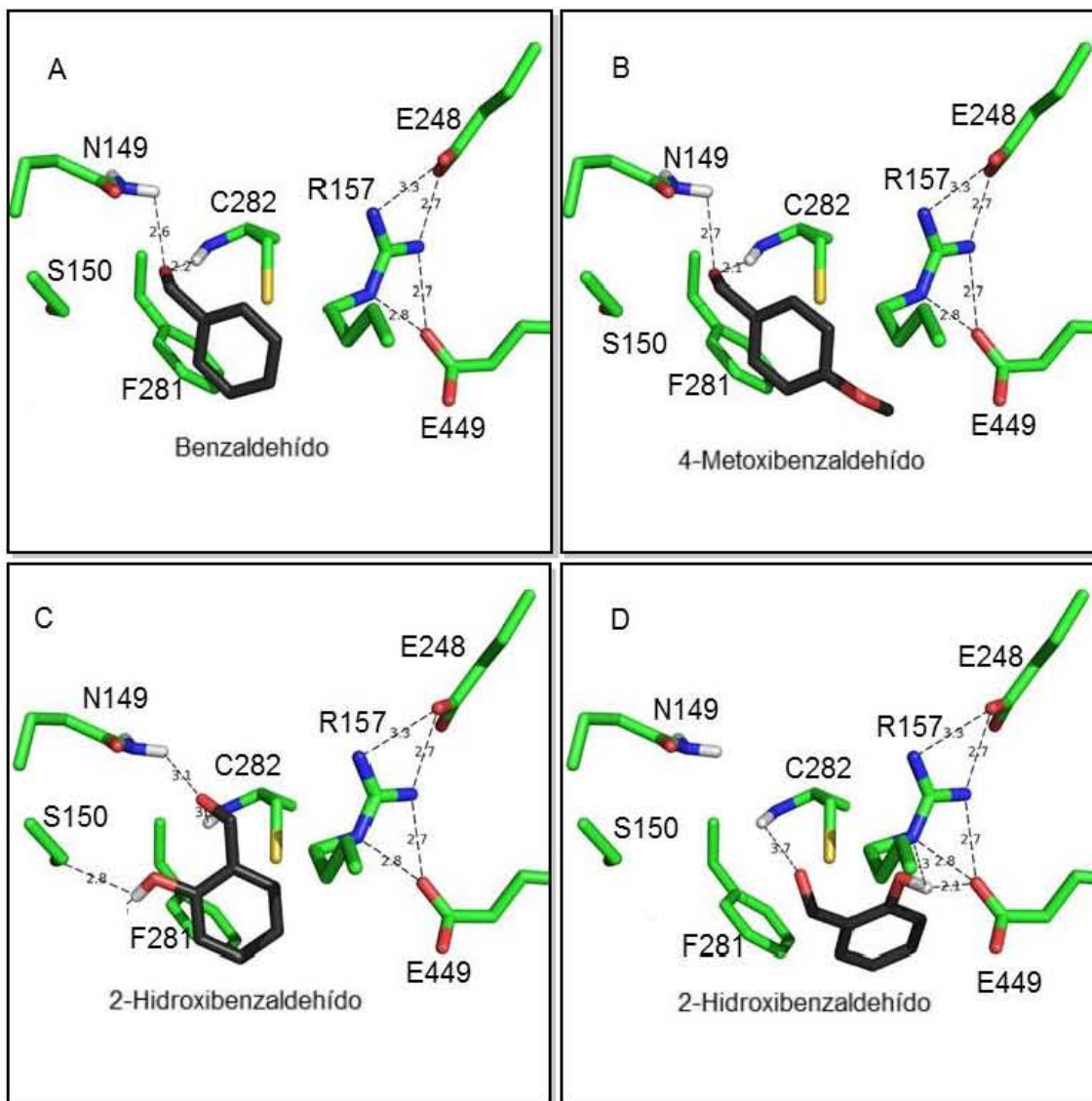
### VI.11 Acoplamiento molecular de los posibles sustratos aldehídos en el sitio activo de la enzima PA2125

Usando el modelo por homología obtenido para PA2125 tratamos de establecer las características estructurales de la enzima que determinan que unos aldehídos sean sustratos

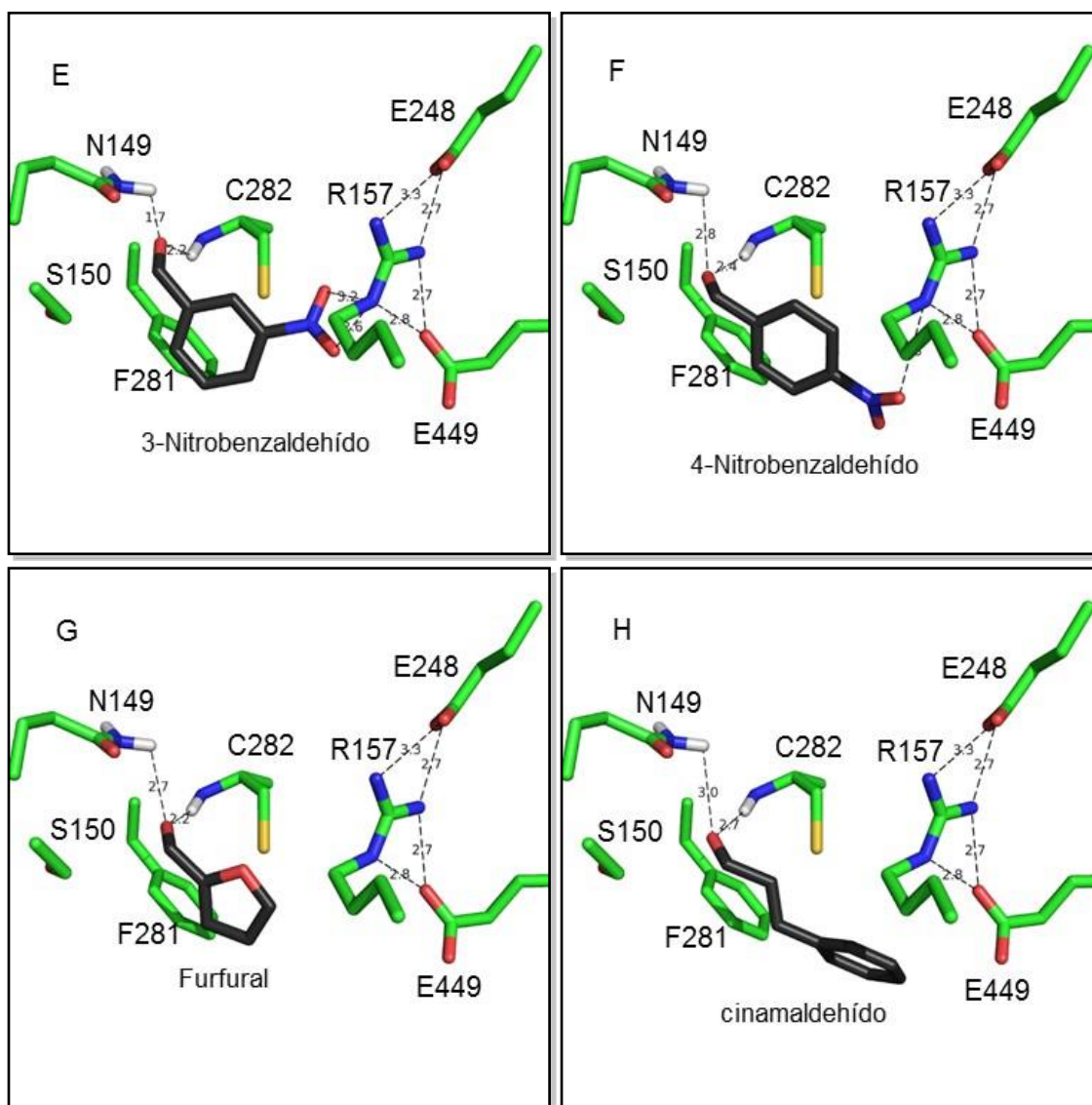
y otros no, y de esta forma explicar los resultados de cinética enzimática. Para ello, se hicieron acoplamientos moleculares con los diferentes aldehídos probados en esta tesis. De los resultados obtenidos, se seleccionaron aquellos modos de unión que eran consistentes con lo reportado en la literatura sobre el mecanismo químico de las ALDHs (revisado en Muñoz-Clares *et al.*, 2010 y Muñoz-Clares *et al.*, 2011) en el que el oxígeno del aldehído debe estar a distancia de puente de hidrógeno del nitrógeno amídico de la cadena principal de la Cys282 y del nitrógeno amídico de la cadena lateral de la Asn149 catalítica (en ambos residuos, numeración de la PA2125), los cuales forman la cavidad del oxianión. Así mismo, el carbono del carbonilo debe estar cercano al azufre del grupo tiol de la cisteína catalítica para que se pueda realizar el ataque nucleofílico por parte de este último. En todos los casos de acoplamiento molecular se cumplieron estos requisitos. Los resultados de los acoplamientos moleculares con cada uno de los aldehídos se encuentran en el Anexo 7.

En los acoplamientos moleculares realizados para los aldehídos aromáticos que sí presentaron actividad como fueron BZAL (Figura 24A), 4-METOXIBZAL (Figura 21B), SAAL (Figura 21C y D), 3-NITROBZAL (Figura 24E), 4-NITROBZAL (Figura 24F), furfural (Figura 24G) cinamaldehído (Figura 24H) encontramos que todos ellos se estabilizan en el sitio activo de PA2125 por medio de interacciones de tipo  $\pi$ -stacking entre el anillo aromático del aldehído y la cadena lateral de Phe281.

Se analizaron dos modos de unión del SAAL en el sitio activo de PA2125. El SAAL, además de las interacciones antes mencionadas, podría estar unido por un puente de hidrógeno entre el grupo OH del anillo aromático del SAAL y la cadena lateral de la Ser150 (Figura 24C). Otra forma de unión del SAAL en el sitio activo de PA2125 permitiría la formación de un puente de hidrógeno entre el grupo -OH del anillo aromático con la cadena lateral del Glu449 (Figura 24D).







**Figura 24.** Acoplamiento molecular (“docking”) de aldehídos aromáticos en el sitio activo de la enzima PA2125. (A) BZAL (modo de unión 7), (B) 4-METOXIBZAL (modo de unión 2), (C) y (D) SAAL (modos de unión 3 y 4), (E) 3-NITROBZAL (modo de unión 8), (F) 4-NITROBZAL (modo de unión 1), (G) furfural (modo de unión 5) y (H) cinamaldehído (modo de unión 1). Se muestra la posición más favorable para cada aldehído dada por el programa y los residuos posiblemente involucrados en la unión de estos aldehídos en el sitio activo. Además, se muestran los puentes de hidrógeno (líneas punteadas negras) entre el oxígeno del grupo carbonilo con el nitrógeno amídico de la cadena lateral de la Asn149 y el nitrógeno amídico de la cadena principal de la Cys282

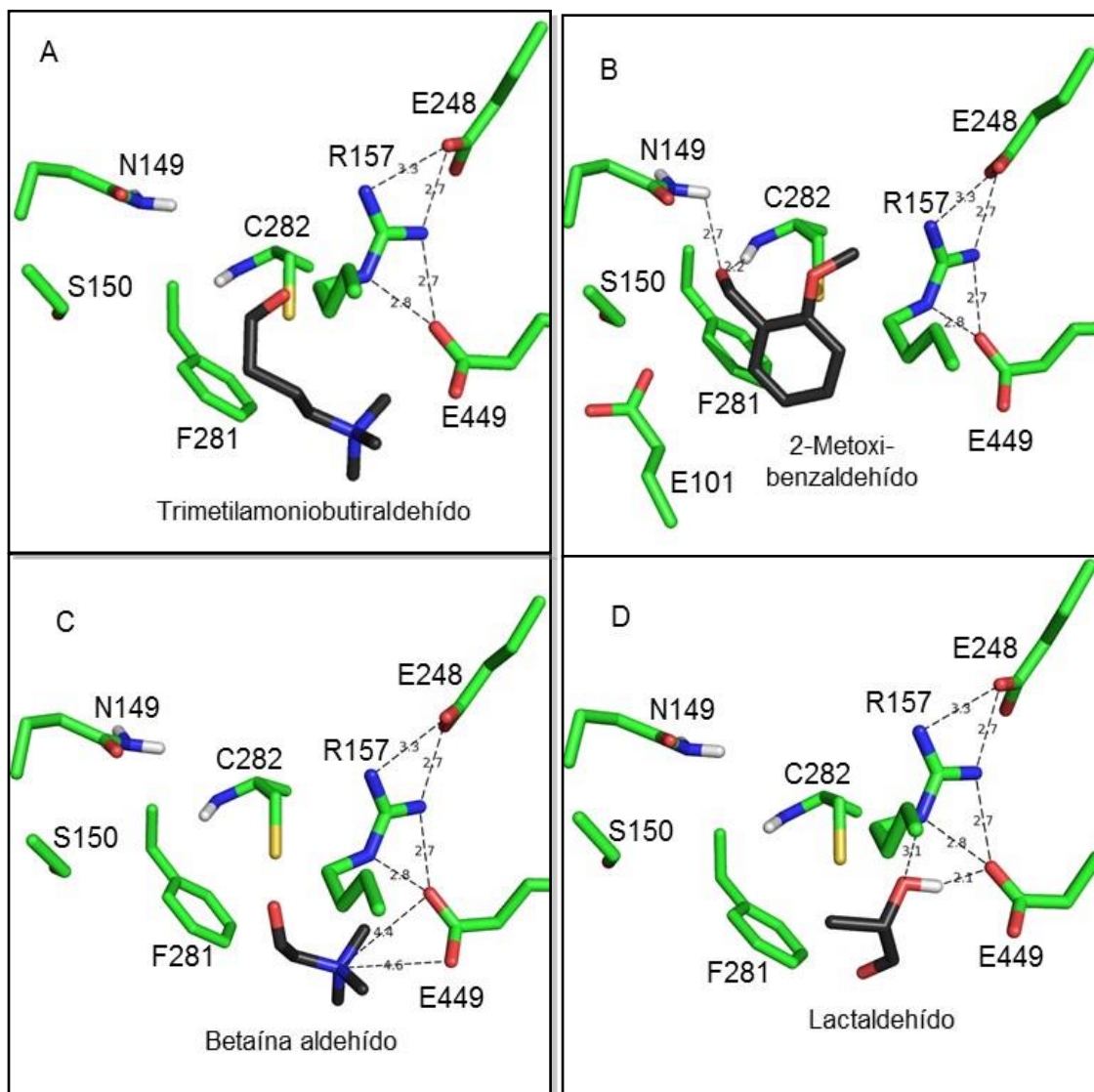
catalíticas. El aldehído se muestra en varillas coloreada por tipo de átomo (color de carbonos en negro, nitrógenos en azul, oxígenos en rojo, hidrógenos en blanco). Para los residuos de aminoácidos de la enzima PA2125 los carbonos se muestran en verde, nitrógenos en azul, oxígenos en rojo, hidrógenos en blanco y azufres en amarillo. Las distancias de los puentes de hidrógeno están en angstroms.

El acoplamiento de 3-NITROBZAL (Figura 24E) indica que el grupo nitro por estar en la posición *meta* con respecto al carbonilo en el anillo aromático interactúa con la cadena lateral de Arg157 atrayéndose electrostáticamente, lo cual contribuye a la unión del aldehído en el sitio activo.

En el caso del 4-NITROBZAL (Figura 24F) pareciera que el grupo NO<sub>2</sub> también interactuara con la cadena lateral de la Arg157.

El furfural tiene un anillo aromático más pequeño que el del BZAL, sin embargo, se puede observar en el acoplamiento molecular (Figura 24G) las interacciones de tipo  $\pi$ -stacking entre el anillo aromático del furfural y la cadena lateral del residuo Phe281.

El acoplamiento con cinamaldehído (Figura 24H) muestra que su anillo aromático estaría demasiado alejado de la cadena lateral del residuo F281 para establecer interacciones de tipo  $\pi$ -stacking. No obstante se puede ver del acoplamiento que se podría unir muy bien a la cisteína catalítica (Cys282).



**Figura 25.** Acoplamiento molecular (“docking”) de TMABAL y 2-hidroxibenzaldehído en el sitio activo de la enzima PA2125. (A) TMABAL (modo de unión 3), (B) 2-hidroxibenzaldehído (modo de unión 10), (C) BAL (modo de unión 3) y (D) lactaldehído (modo de unión 1). Se muestran las posiciones más favorables dadas por el programa y los residuos posiblemente involucrados en la unión de estos aldehídos en el sitio activo. Además, se muestran los puentes de hidrógeno (líneas punteadas negras) entre el oxígeno del grupo carbonilo con el nitrógeno amídico de la cadena lateral de la Asn149 y el nitrógeno amídico de la cadena principal de la Cys282 catalíticas. También se muestran interacciones electrostáticas entre grupo trimetilamonio (de BAL y TMABAL) y los oxígenos

del grupo carboxilo de la cadena lateral del Glu449, así como los puentes de hidrógeno entre grupo 2-OH de lactaldehído y las cadenas laterales de la Arg157 y el Glu449. El aldehído se muestra en varillas coloreada por tipo de átomo (color de carbonos en negro, nitrógenos en azul, oxígenos en rojo, hidrógenos en blanco). Para los residuos de aminoácidos de la enzima PA2125 los carbonos se muestran en verde, nitrógenos en azul, oxígenos en rojo, hidrógenos en blanco y azufres en amarillo. Las distancias están en angstroms.

La enzima PA2125 no presentó actividad con propanal, lactaldehído, BAL, APAL, ABAL y TMABAL. Se llevó a cabo acoplamiento entre el modelo por homología de PA2125 y el TMABAL (Figura 25A), 2-METOXIBZAL (Figura 25B;), BAL (Figura 25C) y lactaldehído (Figura 25D), para tratar de determinar las causas posibles del por qué la enzima no presenta actividad con estos sustratos. Se presume de acuerdo con estos acoplamiento que tanto el TMABAL como betaína aldehído (BAL) interactúan fuertemente en el sitio activo atrayéndose a través de fuerzas electrostáticas entre el grupo trimetilamonio del aldehído y la cadena lateral del residuo Glu449, pero a su vez existirían repulsiones electrostáticas entre el grupo trimetilamonio y la cadena lateral de la Arg157, como se puede ver en el acoplamiento (Figuras 25A y 25C) lo cual explicaría porque PA2125 no presenta actividad con estos compuestos.. Se puede observar en el caso de 2-METOXIBZAL (Figura 25B) en el que el oxígeno carbonílico del compuesto interactúa tanto con la Asn149 como con la cisteína, por lo cual el modelo y los acoplamiento no permiten dar explicación al hecho de que la enzima no presente actividad con estos compuestos.. Finalmente, el acoplamiento con lactaldehído mostró que el grupo 2-OH interactúa a través de puentes de hidrógeno con el oxígeno del grupo carbonilo de la cadena lateral del Glu449 y con la cadena lateral de la Arg157, razón por la cual tampoco permite que el carbono carbonílico del aldehído llegue a acercarse a la cisteína catalítica como para formar

un enlace covalente y esto si nos ayudaría a explicar el hecho de que la enzima *PA2125* no presentó actividad con lactaldehído.

## VII DISCUSIÓN DE RESULTADOS

### VII.1 Papel fisiológico de la *PA2125*

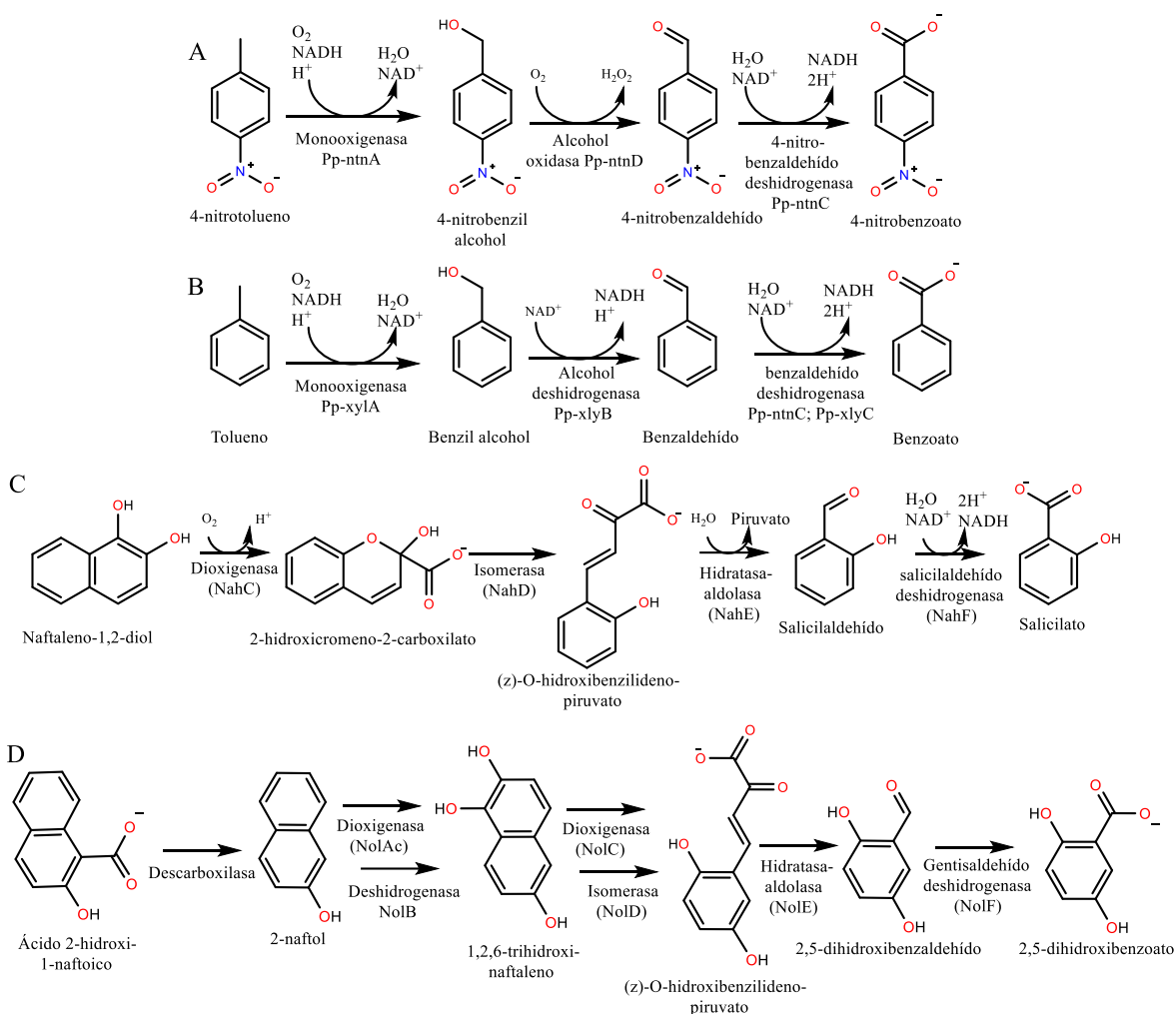
La enzima *PA2125* es una ALDH que no había sido caracterizada hasta la fecha, ni funcional ni estructuralmente, aunque en el genoma de *P. aeruginosa* el gen *pa2125* había sido anotado como que codificaba para una BADH, lo que sugería que *PA2125* pudiese oxidar a BAL formando glicina betaína y participar en la degradación de colina, al igual que hace la *PaBADH* (*PA5373*) supliendo a ésta en el caso de estar inactiva. La *PaBADH* está codificada por el gen *betB* que forma parte del operón *betIBA* junto con el gen *betA* (que codifica a una colina deshidrogenasa) y el gen *betI* (que codifica a un represor transcripcional). La colina deshidrogenasa de *P. aeruginosa* (*PA5372*) oxida colina para producir BAL, aldehído que es oxidado por la *PaBADH* para producir glicina betaína. La colina deshidrogenasa pertenece a la familia de proteínas de dominio conservado PRK02106, que a su vez pertenece a la superfamilia glucosa-metanol-colina oxidorreductasas C (GMC oxidorreductasas C) de dominio conservado cl27052. En el contexto genómico del gen *pa2125* (Figura 24) se encuentra el gen *pa2124* que codifica a una deshidrogenasa perteneciente a la superfamilia GMC oxidorreductasas C. Ambos genes, *pa2124* y *pa2125*, parecen formar un operón y es de suponerse que la enzima *PA2124* oxide a un alcohol y produzca el aldehído que es sustrato de la enzima *PA2125*. Esto era un indicio más para suponer que las enzimas *PA2124* y *PA2125* podrían suplir a

las enzimas de *P. aeruginosa* colina deshidrogenasa (*PA5372*) y BADH (*PA5373*; *PaBADH*), respectivamente, y que estuvieran participando en una ruta catabólica alterna de colina o de síntesis de glicina betaína, aunque no es muy frecuente que las bacterias posean dos rutas para degradar el mismo compuesto. Por medio de alineamientos múltiples encontramos que *PA2125* presenta una identidad del 34% con *PA5373* (*PaBADH*) y *PA2124* tiene una identidad del 38% con *PA5372* (colina deshidrogenasa), lo que indica que estas enzimas no son de la misma familia, aunque sí sean de la misma superfamilia, y que por tanto sus sustratos podrían ser diferentes y participarían en rutas metabólicas diferentes. Esto se apoya, además, en que los residuos determinantes de especificidad de las enzimas *PA2125* y *PA5373* son muy diferentes. La conclusión de que *PA2125* no está participando en la ruta de degradación de colina ni en la síntesis de glicina betaína se confirmó experimentalmente porque *PA2125* no presentó actividad con BAL y sí con aldehídos aromáticos derivados del BZAL.

Por otro lado, en el contexto genómico de *PA2125* hay otros dos genes que codifican proteínas que podrían participar en una ruta metabólica junto a *PA2124* y *PA2125*. Una de estas proteínas es *PA2122*, la que de acuerdo con el CDD NCBI pertenece a una familia, cuyos miembros presentan las actividades poliquétido ciclasa/deshidrasa y transporte de lípidos, dentro de la superfamilia cl14643. Esta superfamilia comprende proteínas que presentan un sitio de unión a un ligando hidrofóbico lo que concuerda con la preferencia de la enzima *PA2125* por aldehídos aromáticos encontrada en este trabajo. La enzima *PA2122* podría formar el compuesto aromático hidroxilado que usa *PA2124* para oxidarlo al aldehído sustrato de la ALDH *PA2125*.

Por lo anterior, en la base de datos MetaCyc (Caspi *et al.*, 2017) buscamos rutas metabólicas degradativas de bacterias en las que se produjeran como intermediarios los

aldehídos aromáticos que habíamos encontrado que eran sustratos de la PA2125 *in vitro* y en las que, por tanto, esta enzima pudiera participar; estas rutas se muestran en el Esquema 6. No se encontró ninguna para el 3-NITROBZAL ni siquiera una enzima anotada como 3-NITROBZAL deshidrogenasa.

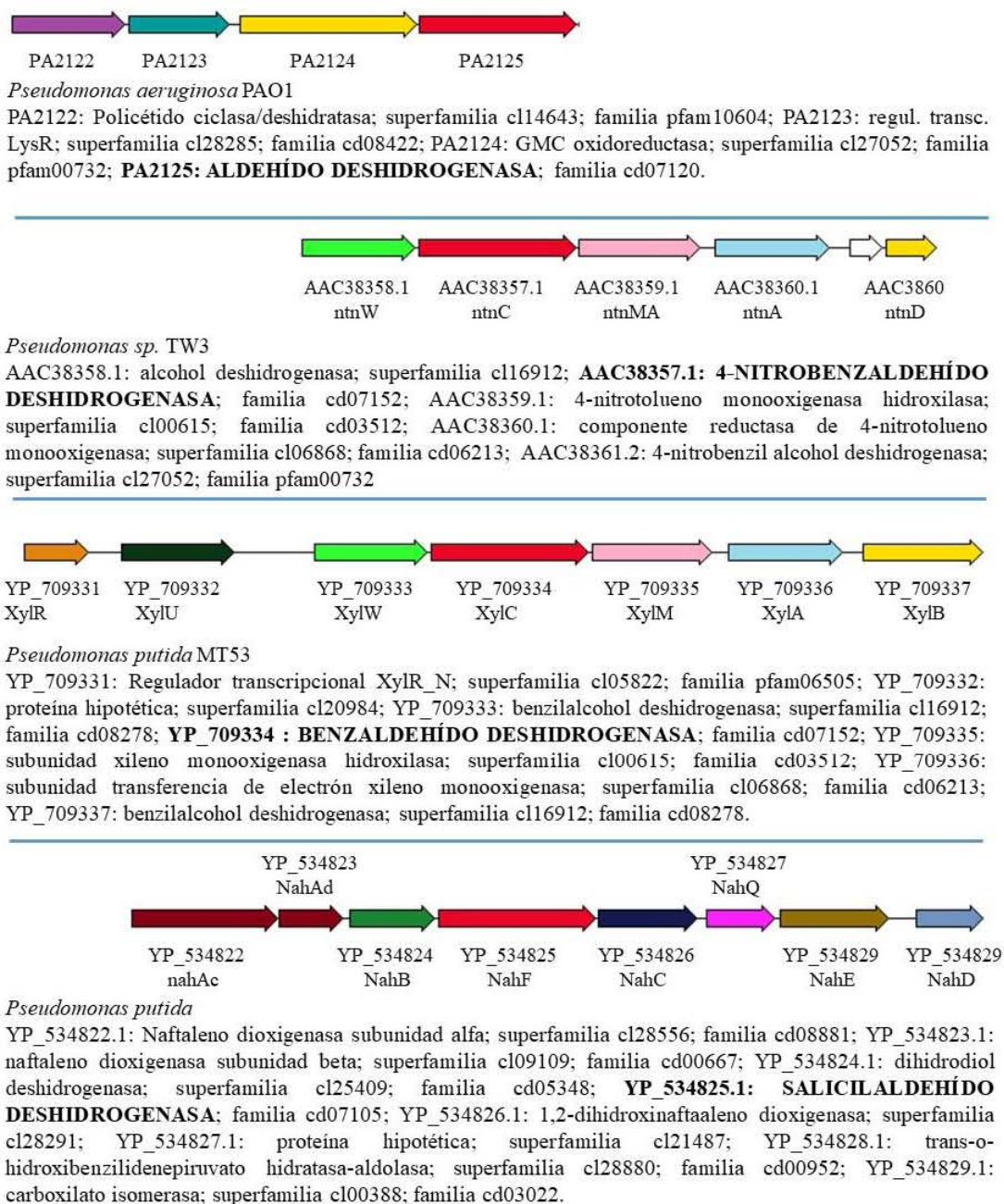


**Esquema 6. Rutas metabólicas en las que participan los aldehídos que *in vitro* son sustratos de la PA2125.** (A) Ruta de degradación de 4-nitrotolueno en *P. putida* (James y Williams, 1998; James *et al.*, 2000); (B) Ruta de degradación de tolueno a benzoato en *P. putida* (Shaw y Harayama, 1990; Haigler y Spain, 1993; Inoue *et al.*, 1995); (C). Ruta de degradación de naftaleno en *P. putida* y *P. sp.* (Barnsley, 1976); (D) Ruta de degradación de naftaleno y ácido 2-hidroxi-1-naftoico vía ácido gentísico en *Burkholderi sp.* BC1 (Chowdhury *et al.*, 2014).

Buscamos también los contextos genómicos de genes que codifican a las ALDHs que participan en estas rutas metabólicas y se compararon con el del gen que codifica a la PA2125. En el Esquema 6 y en la Figura 24 se puede observar que las rutas de degradación de 4-nitrobenceno y de benceno son iguales, así como los contextos genómicos, lo que sugiere que la misma ALDH puede oxidar los dos aldehídos que se producen en estas rutas (el 4-NITROBZAL y el BZAL). Por otro lado, el contexto genómico de los genes que participan en la ruta de degradación de naftaleno, en la que se forma salicilaldehído (2-hidroxibenzaldehído) es muy parecido al de la posible ruta de degradación en la que se produce gentisalaldehído (2,5-hidroxibenzaldehído) (ver la Figura 8 de la sección de Resultados). Si bien las ALDHs de estas dos rutas pertenecen a familias diferentes pareciera que ambas enzimas podrían usar a los dos sustratos. Hay que mencionar que los genes de las rutas de degradación de 4-nitrobenceno, benceno y tolueno están en plásmidos de la bacteria *P. putida*, mientras que los de la ruta de degradación de naftaleno están en el DNA genómico de esta bacteria. Debido a que el genoma de *P. putida* también posee el gen de la una enzima ortóloga a PA2125 (con la que tiene un 99% de identidad) y está en un contexto genómico idéntico, concluimos que esta última enzima no es ni una salicilaldehído deshidrogenasa ni una gentisalaldehído deshidrogenasa, sino que tendrá el mismo sustrato fisiológico que la PA2125, sustrato aún desconocido. Esta conclusión además se apoya en lo que se discute a continuación. En la familia ALDH cd07120 a la que pertenece PA2125 existe una enzima ALDH de *P. putida* codificada por el gen *psfA*, la cual se le indujo su transcripción en presencia de furfural alcohol, furfural y ácido furoico, por lo cual se le presumió actividad furfural deshidrogenasa (Nichols y Mertens, 2008). Nuestros resultados de actividad con furfural como sustrato y de los análisis de conservación de residuos del



sitio activo y de contextos genómicos indican que el furfural no es el sustrato fisiológico de la PA2125.



**Figura 26. Comparación de contextos genómicos de PA2125 con otras ALDHs cuyo sustratos lo son también de PA2125 . En rojo el gen que codifica la aldehído deshidrogenasa: PA2125 (aldehído deshidrogenasa); Benzaldehído**

deshidrogenasa (YP\_709334); 4-nitrobenzaldehído (AAC38357); salicilaldehído  
deshidrogenasa (YP\_534826)

El gen que codifica para la enzima PA2125 está presente en todos los genomas de las cepas de *P. aeruginosa* secuenciados hasta la fecha (99), tanto patógenas, por ejemplo la DK2 que fue aislada de pacientes con fibrosis quística (Rau *et al.*, 2012), como no patógenas. Por ello pensamos que esta enzima no está relacionada con la virulencia de la cepa, ya que los genes necesarios para la virulencia se han perdido en las especies no patógenas (Bianconi *et al.*, 2015). Esto es además consistente con que es altamente improbable que la bacteria encuentre el tipo de aldehído aromático que muy probablemente es su sustrato en los tejidos que infecta. Enzimas con un porcentaje de identidad del 99% con PA2125 están presentes en cinco especies de proteobacterias (Anexo 8), todas ellas patógenas oportunistas pero además pertenecientes a géneros que poseen especies capaces de crecer en suelos: *Acinetobacter baumannii*; *Enterobacter cloacae*; y *Pseudomonas putida*, con gran potencial de degradación de compuestos aromáticos. En todas estas bacterias, el contexto genómico es el mismo por lo que las enzimas deben participar en la misma ruta metabólica, aunque ninguna de ellas ha sido caracterizada hasta la fecha. Con menor identidad, pero superior al 60 %, hay enzimas ortólogas en otras especies de *Pseudomonas*, *Burkholderia*, entre ellas *B. cepacia* que presenta una gran capacidad degradativa de contaminantes orgánicos; de *Acinetobacter*, de *Achromobacter*, de *Pseudonocardia*, de *Agrobacterium* y de *Stenotrophomonas*. Todas ellas poseen en su genoma al menos los dos genes ortólogos de PA2124 y PA2125, por lo que pensamos que deben poseer la ruta metabólica en la que participa PA2125. De nuevo, todas estas bacterias son capaces también de crecer en suelos. Por todo lo anterior, pensamos que lo más probable es que la

*PA2125* esté involucrada en la ruta degradativa de un compuesto aromático que puede ser un contaminante de suelos y por tanto tener potencial biotecnológico en biorremediación. Los compuestos aromáticos se encuentran en el ambiente como contaminantes provenientes del petróleo y las gasolinas. Varios de ellos como el BZAL y algunos de sus derivados son compuestos xenobióticos tóxicos presentes en una gran cantidad de efluentes de la industria farmacéutica, química y petroquímica (Capestany *et al.*, 1997).

El hecho de que *PA2125* y gentsaldehído deshidrogenasa (acceso NCBI: APP18125.1) tienen igual los residuos Arg157 y Glu449 (posición en *PA2125*), y residuos pequeños en la posición Ser150 (posición en *PA2125*) que en la gentsaldehído deshidrogenasa es una alanina, afianza el hecho de que *PA2125* utilice como sustrato al gentsaldehído. Se llevó a cabo un acoplamiento molecular con gentsaldehído en el sitio activo del modelo por homología de *PA2125* para determinar cuáles residuos puedan participar de la unión del aldehído y se encontró que posiblemente se pueden dar las mismas interacciones entre el grupo 2-OH del anillo aromático del gentsaldehído (2,5-dihidroxibenzaldehído) y la cadena lateral del residuo Ser150, como las que encontramos cuando hicimos el acoplamiento molecular con SAAL. Se podría dar un puente de hidrógeno entre el grupo 5-OH del anillo aromático y un oxígeno del grupo carboxilo del Glu449, sin embargo, la distancia no corresponde a un puente de hidrógeno pero hay que recordar que se trabaja con un modelo por homología y hay que aceptar cierto error en la medida.

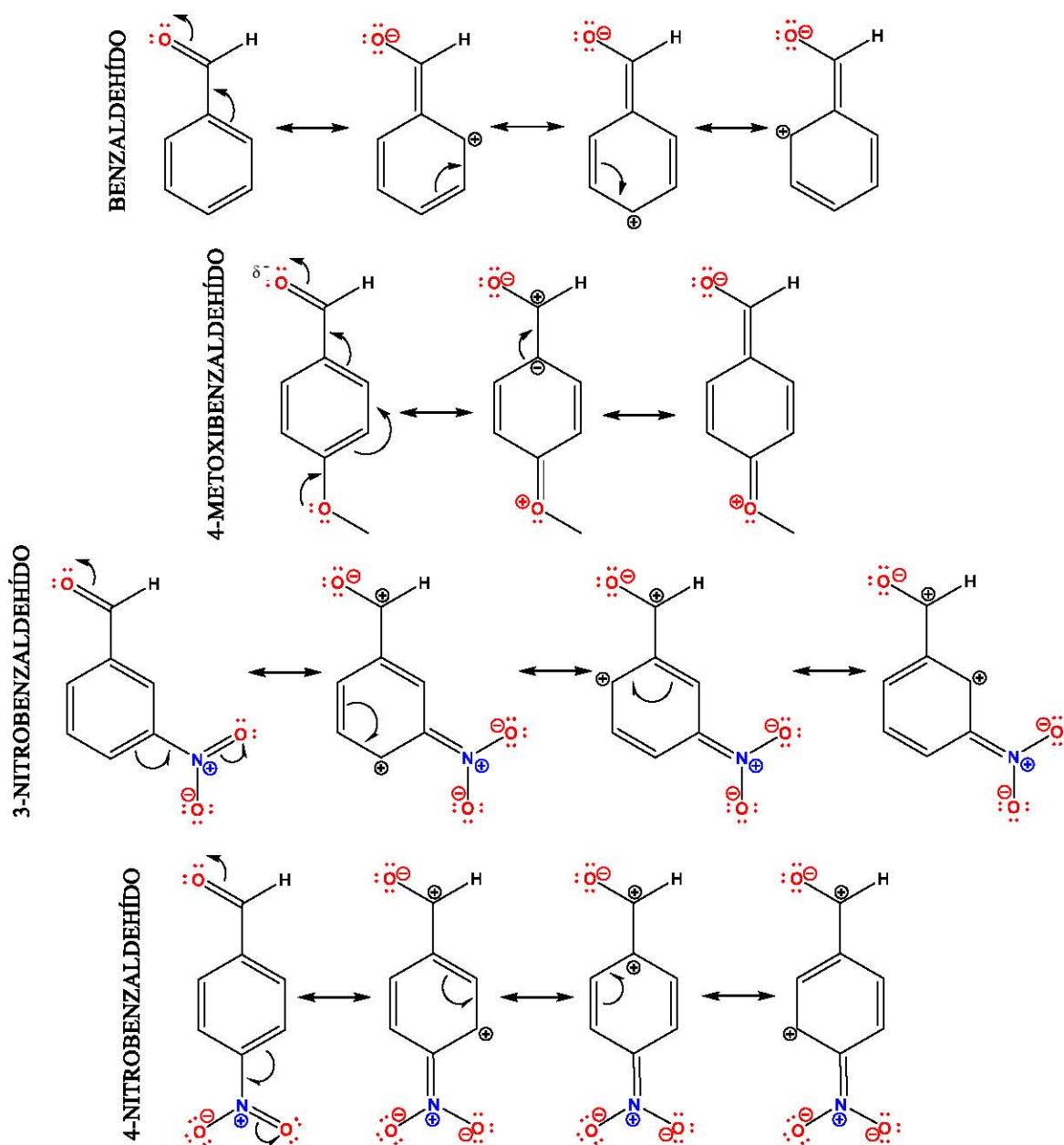
Por todo lo anterior pensamos que *PA2125* puede unir a aldehídos aromáticos disustituidos por grupos hidroxilo o grupos sustituyentes polares en las 2 y 3 o 2 y 5 del anillo aromático, aldehídos que se pueden producir como intermediarios en rutas metabólicas de degradación. Los residuos Ser150, Arg157 y Glu449 parece que participan en la unión de aldehídos anteriormente descritos al sitio activo de la enzima *PA2125*. La familia ALDH

cd07120 presentan un residuo pequeño en la posición 150 (numeración de *PA2125*) sea una serina o una alanina, esto debido posiblemente a que la presencia de un residuo grande como fenilalanina o tirosina (como en otras familias ALDHs) impida estéricamente el acercamiento del aldehído a la cisteína catalítica o reduzca el espacio para el acomodamiento del aldehído. Parece que el Glu449 juega un papel importante en la selectividad de la enzima por sustratos ya que como se vio en los acoplamientos moleculares este residuo interacciona con grupos amonio, amino o grupos hidroxilo permitiendo o no el acercamiento del grupo carbonilo a hueco del oxianión.

## **VII.2 Mecanismo catalítico de la *PA2125***

Los grupos sustituyentes en un anillo aromático pueden presentar tanto efectos inductivos como de resonancia, como es el caso del grupo carbonilo que presenta un fuerte efecto inductivo atractor hacia el oxígeno y a la vez fuertes efectos de resonancia entre el carbono carbonílico y el carbono del anillo aromático. Los sustituyentes halógeno, hidroxilo, alcoholilo y amino tienen efectos inductivos atractores de electrones debido a la electronegatividad del átomo -X, -O o -N unido al anillo aromático, pero tienen efectos de resonancia donadores de electrones debido al par de electrones no enlazado en los mismos átomos -X, -O o -N. Cuando los dos efectos actúan en direcciones opuestas, domina el más fuerte de los dos; por ejemplo, los halógenos presentan pares de electrones no enlazantes que les permite efectos de resonancia pero son muy electronegativos, por lo tanto el efecto atractor de electrones prima sobre los efectos donadores por resonancia, mientras que en un grupo amino predominan los efectos donadores de electrones por resonancia. Los grupos carbonilo, ciano y nitro presentan efectos inductivos y de resonancia atractores de

electrones. El BZAL es mucho menos reactivo que los aldehídos alifáticos en las reacciones de adición nucleofílica, esto se debe al efecto electrodonador por resonancia del anillo aromático, lo que hace que el carbono carbonílico sea menos electrofílico (Esquema 7A).



*Esquema 7. Estructuras resonantes de aldehídos aromáticos con grupos electrodonadores y electroatrayentes.*

Los grupos sustituyentes que pueden estar unidos al anillo aromático del BZAL pueden aumentar o disminuir la electrofilicidad del carbono carbonílico. Los grupos electroatrayentes como el grupo  $-\text{NO}_2$  en la posición *para* (4- $\text{NO}_2$ ) (Esquema 7B) y en la posición *meta* (3- $\text{NO}_2$ ) (Esquema 7C) aumentan su electrofilicidad debido a que producen deficiencia electrónica por efecto de resonancia, haciéndolo más reactivo a un ataque nucleofílico. Por el contrario, los grupos sustituyentes electrodonadores, como el grupo  $-\text{OCH}_3$  (Esquema 7D), aumentan la densidad electrónica cuando están en las posiciones *meta* y *para*, igualmente por efecto de resonancia, de tal manera que el carbono carbonílico disminuye su reactividad frente a un ataque nucleofílico al disminuir su electrofilicidad. La ecuación de Hammett (Ecuación 9) y las constantes de sustitución de Hammett permiten cuantificar estos efectos. Las constantes de sustitución que miden el poder de sustitución o donación de electrones en un anillo de benceno se derivan empíricamente a partir de los  $\text{pK}_a$ s de ácidos benzoicos sustituidos:

$$(\text{pK}_a)_x = (\text{pK}_a)_0 - \sigma_x \quad \text{Ecuación 9}$$

Donde X es un sustituyente en la posición *meta* o *para* del ácido benzoico,  $(\text{pK}_a)_x$  es el  $\text{pK}_a$  del ácido sustituido,  $(\text{pK}_a)_0$  es el  $\text{pK}_a$  del ácido sin sustituyente y  $\sigma_x$  es la constante de sustitución para X. En la Tabla 8 se incluyen las constantes de sustitución de algunos grupos electroatrayentes y electrodonadores en las posiciones *para* y *meta* del ácido benzoico.

**Tabla 8. Constantes de sustitución para y meta obtenidas de la ecuación de Hammett para grupos sustituyentes en el anillo de benzeno.** La constante de sustitución ( $\sigma$ ) sólo se mide para los grupos sustituyentes en posición para y meta. Cuanto más positivo sea el valor mayor será la deficiencia electrónica del carbono del grupo carboxilo del ácido benzoico (Ritchie y Sager, 1964).

Tipo de sustituyente	Sustituyente	Efecto para	Efecto meta
Electrodonador	Dimetilamino	-0.83	-0.211
	Amino	-0.66	-0.161
	Metoxi	-0.268	0.115
	Etoxi	-0.25	0.015
	Metil	-0.170	-0.069
Electroatrayente	Fluoro	0.062	0.337
	Cloro	0.227	0.373
	Bromo	0.232	0.339
	Yodo	0.276	0.353
	Ciano	0.66	0.56
	Nitro	0.778	0.710

Por lo anterior, y teniendo en cuenta el mecanismo de reacción de la ALDHs, esperaríamos que un grupo 3-NO<sub>2</sub> o 4-NO<sub>2</sub> en el anillo de BZAL hiciera al carbono carbonílico más electrofílico y más susceptible al ataque nucleofílico por parte del tiol de la cisteína catalítica en la etapa de acilación, y también harían más electrofílico al carbono del intermediario tioéster haciéndolo más susceptible al ataque nucleofílico del agua catalítica para producir el producto ácido en la etapa de desacilación. En cambio el grupo 4-OCH<sub>3</sub> disminuiría la electrofilicidad del carbono en el grupo carbonilo y en el intermediario tioéster haciéndolo menos susceptible a un ataque nucleofílico.

$$(pK_a)_x = (pK_a)_0 - \sigma_x$$

Ecuación 10

Donde X es un sustituyente en la posición *meta* o *para* del ácido benzoico,  $(pK_a)_x$  es el  $pK_a$  del ácido sustituido,  $(pK_a)_0$  el del ácido sin sustituyente y  $\sigma_x$  es la constante de sustitución para X. En la Tabla 8 se pueden observar algunos grupos electroatrayentes y electrodonadores con sus respectivas constantes de sustitución para las posiciones *para* y *meta* en el ácido benzoico. Por lo anterior y teniendo en cuenta el mecanismo de reacción de la ALDHs vemos que un grupo 3-NO<sub>2</sub> o 4-NO<sub>2</sub> en el anillo de BZAL permite que el carbono carbonílico sea más electrofílico y más susceptible al ataque nucleofílico por parte del tiol de la cisteína catalítica en la etapa de acilación, y a la vez harían más electrofílico al carbono del intermediario tioéster haciéndolo más susceptible al ataque nucleofílico del agua catalítica para producir el producto ácido en la etapa de desacilación. En cambio el grupo 4-OCH<sub>3</sub> disminuye la electrofilicidad del carbono en el grupo carbonilo y en el intermediario tioéster haciéndolo menos susceptible a un ataque nucleofílico.

De acuerdo con los resultados de la caracterización cinética obtenidos en este trabajo de investigación, los sustratos preferidos por la enzima PA2125 a las concentraciones fisiológicas de NAD<sup>+</sup> en la bacteria fueron el 3-NITROBZAL y el 4-NITROBZAL, ya que presentaron las mayores constantes de especificidad aparentes ( $apk_{cat}/K_m$ ) y una mayor afinidad (medida como  $K_m$ ), aunque la velocidad máxima es menor con 4-NITROBZAL. Quisimos saber cuáles eran las características que hacían a estos dos compuestos mejores sustratos para PA2125 que el BZAL, para lo cual realizamos ensayos de cinética rápida para determinar cuál es el paso del mecanismo catalítico limitante de la reacción para cada aldehído y cómo se afecta la velocidad de las dos etapas del mecanismo catalítico por la presencia en el anillo aromático de grupos sustituyentes electrodonadores o electroatrayentes. Para este análisis tomamos al BZAL como referencia. El BZAL presenta una amplitud del estallido ( $A_{burst}$ ) muy cercana a la concentración de sitios activos de PA2125, por lo cual podemos



decir que la constante  $k_1'$  (constante de velocidad de pseudoprimer orden de acilación)  $\gg k_3$  (constante de velocidad de primer orden de desacilación) y por tanto la constante de velocidad observada para el estallido ( $k_{burst}$ ) es igual a  $k_1'$ . A su vez nos indica que la etapa lenta o limitante de la velocidad de la reacción para el BZAL es la desacilación. El 4-NITROBZAL, que posee un grupo sustituyente electroatrayente en la posición *para* del anillo aromático del BZAL, mostró valores muy cercanos tanto de la amplitud del estallido como de la  $k_{burst}$  a los estimados para BZAL y una velocidad de estado estacionario menor para el 4-NITROBZAL que para el BZAL. Este último resultado concuerda con los datos obtenidos en las medidas de estado estacionario; sin embargo no es consistente con las características del grupo electroatrayente nitro, que en la posición *para* del anillo aromático debería hacer más electrofílico al carbono carbonílico tanto del aldehído como del tioéster y por tanto aumentar la velocidad tanto del paso de acilación como de desacilación. Pensamos que esta comportamiento diferente al esperado se debe a las interacciones que hace el grupo nitro del 4-NITROBZAL en el sitio activo de la enzima, si bien contribuyen a la mejor unión de este sustrato que el BZAL, probablemente coloquen al carbono carbonílico de este compuesto en una posición menos favorable para el ataque nucleofílico por parte de la Cys catalítica, y aún menos de la molécula de agua hidrolítica, que el carbono carbonílico del BZAL y ello tiene como consecuencia que no haya un aumento en la velocidad de acilación y que incluso haya una disminución en la velocidad de desacilación con respecto al BZAL. En el caso del 3-NITROBZAL el efecto del grupo nitro electroatrayente en la posición *meta* del anillo aromático del BZAL es menor que cuando está en la posición *para*, aunque aun así debe hacer más electrofílico al carbono del carbonilo del aldehído y del intermediario tioéster. Vimos sin embargo que la  $k_{burst}$  es mucho menor para el 3-NITROBZAL que para el BZAL, mientras que la velocidad de

estado estacionario,  $v_{ss}$ , es casi igual, también en concordancia con lo observado en los estudios de estado estacionario. Parece por tanto que las interacciones que el grupo 3-NO<sub>2</sub> hace en el sitio activo, importantes para la unión del aldehído, también están modulando el efecto de este sustituyente, en este caso afectando bastante más al paso de acilación que al paso de desacilación lo que se refleja en que las constantes de velocidad de ambos pasos casi se igualan por lo que la amplitud del burst está muy disminuida y es mucho menor que la concentración de sitios activos. Por último, quisimos completar este estudio investigando los efectos sobre las dos etapas de la reacción catalizada del grupo electrodonador -OCH<sub>3</sub> en la posición *para* del anillo aromático del BZAL del 4-METOXIBZAL, grupo que debería hacer menos electrofílico al carbono carbonílico y por lo tanto menos susceptible a un ataque nucleofílico por parte de la cisteína catalítica y de la molécula de agua hidrolítica. Esta predicción concuerda con los valores menores tanto de  $k_{burst}$  como de la  $v_{ss}$  en comparación con los valores del BZAL, y también es consistente con la menor velocidad de reacción que observamos en los estudios de estado estacionario cuando se usó el 4-METOXIBZAL como sustrato. Tomando en consideración los resultados de cinética rápida y de estado estacionario, pensamos que este aldehído es el que peor se une a la enzima y por tanto el sitio activo tiene menos influencia sobre los efectos que el grupo sustituyente ejerce sobre el carbono del carbonilo.

### VII.3 Mecanismo cinético de PA2125

Los patrones de velocidad inicial obtenidos con BZAL y NAD(P)<sup>+</sup> son consistentes con un mecanismo Bi Bi Ordenado en estado estacionario (Cleland, 1963) en el que el NAD<sup>+</sup> es el primer sustrato en unirse a la enzima y el NADH el último producto en salir del sitio activo

(Esquema 4). Este mecanismo se ha reportado para otras ALDHs hidrolíticas (MacGibbon *et al.*, 1977; Henehan y Tipton, 1992; Valenzuela-Soto *et al.*, 1993; Velasco-García *et al.*, 2000) y es consistente con el mecanismo químico de la reacción, pero no es fácil de explicar en términos estructurales dado que la arquitectura del sitio activo de las ALDHs parece permitir la entrada al azar del aldehído y del cofactor. Tampoco podemos por ahora explicar en términos estructurales, como ya se mencionó, las diferencias cinéticas de las reacciones dependientes de  $\text{NAD}^+$  y de  $\text{NADP}^+$ . En la reacción  $\text{NAD}^+$ -dependiente, la enzima muestra una afinidad (medida como  $K_{ia}$ ) 20 veces mayor por el  $\text{NAD}^+$  que por el  $\text{NADP}^+$ , y una  $k_{cat}$  5 veces mayor, mientras que la afinidad por el aldehído (medida como  $K_m$ ) es 3 veces menor. Es claro que la presencia o no del grupo 2'-fosfato en el anillo de la ribosa de la adenosina afecta la manera en que el cofactor (probablemente tanto en su forma oxidada como el reducida) se une a la enzima, y esto a su vez tiene implicaciones en la cinética de la reacción. Los valores menores de  $K_m^{\text{NAD(P)}^+}$  que el valor de  $K_m^{\text{BZAL}}$  podrían indicar que bajo las condiciones normales fisiológicas de la bacteria el primer sustrato en unirse sería el  $\text{NAD}^+$ , pero eso dependerá además de las concentraciones intracelulares de ambos. Además, en las gráficas de dobles recíprocos se observa que las líneas se cruzan en el tercer cuadrante y, por lo tanto, los valores aparentes de  $V$ ,  $K_m$  y  $V/K_m$  cambian cuando se cambia la concentración del aldehído o del cofactor, de manera que la enzima presenta una mayor  $K_m$  aparente por el  $\text{NAD}^+$  cuando hay concentraciones altas del aldehído y viceversa, una menor  $K_m$  aparente cuando hay concentraciones bajas del aldehído (compárese la  $apK_m^{\text{BZAL}}$  obtenida en los experimentos de saturación a 0.6 mM  $\text{NAD}^+$  con la  $K_m^{\text{BZAL}}$  real obtenida por medio de los patrones de velocidad inicial: 2 mM frente a 13.4 mM). Lo anterior parece sugerir que el aldehído está interfiriendo en la unión del  $\text{NAD}^+$  a la enzima y, si ese fuese el caso, una posible explicación sería que el aldehído llegue al sitio activo

por la entrada del nucleótido, lo cual sería más probable cuando haya concentraciones altas del aldehído. Sin embargo, el valor relativo de las constantes de velocidad  $K_m^{\text{NAD}^+}$  y  $K_{ia}$ , cuando  $K_m^{\text{NAD}^+}$  es mayor que  $K_{ia}$ , hacen que se obtenga el patrón de líneas que se cruzan en el tercer cuadrante. Teniendo en cuenta la definición de  $K_m^{\text{NAD}^+}$  y  $K_{ia}$  en términos de constantes de velocidad para un mecanismo Bi Bi ordenado en estado estacionario ( $K_m^{\text{NAD}^+} = k_{+3}k_{+4} / k_{+1}(k_{+3} + k_{+4})$  y  $K_{ia} = k_{-1}/k_{+1}$ ; Segel, 1993) y el hecho de que en las enzimas ALDH la liberación del producto ácido es muy rápida comparada con la liberación del nucleótido reducido (es decir,  $k_{+3}$  es mucho mayor que  $k_{+4}$ ; véase el Esquema 4), la condición de que  $K_m^{\text{NAD}^+}$  sea mayor que  $K_{ia}$  se da cuando la liberación del NADH es más rápida que la liberación del  $\text{NAD}^+$ , es decir, cuando  $k_{+4}$  es mayor que  $k_{-1}$ . Este comportamiento se ha observado en los mecanismos cinéticos descritos para las ALDHs de hígado de humano (Sidhu y Blair, 1975), la BADH de *P. aeruginosa* (Velasco-García *et al.*, 2000) y la KauB de *P. aeruginosa* (Cardona-Cardona, 2017). Otras ALDHs muestran un patrón de velocidad inicial de líneas que se cruzan sobre el eje, indicando que la  $K_{ia}$  y la  $K_m^{\text{NAD}^+}$  son iguales (Valenzuela-Soto y Muñoz-Clares, 1993; Figueroa-Soto y Valenzuela-Soto, 2000). El tener un valor de  $K_{ia}$  mucho menor que el de  $K_m^{\text{NAD}^+}$  explica el aparente mecanismo Ping Pong reportado para otras ALDHs (Mori *et al.*, 1980; Falkenberg y Strøm, 1990), mecanismo cinético que no es compatible con el mecanismo químico de estas enzimas. En el caso de la PA2125 la  $K_{ia}^{\text{NAD}^+}$  es 40 veces menor que la  $K_m^{\text{NAD}^+}$ , pero la  $K_{ia}^{\text{NADP}^+}$  es sólo 2.5 menor que la  $K_m^{\text{NADP}^+}$ , lo que indica que si bien la afinidad de la enzima libre por el  $\text{NADP}^+$  es mucho menor que por el  $\text{NAD}^+$ , unas veinte veces menor, las diferencias en la afinidad entre el NADH y el NADPH deben ser mucho menores. Las razones estructurales de estas características cinéticas de la enzima no las conocemos por el momento.

Un aspecto de estas propiedades del mecanismo cinético de la enzima que pudiera tener relevancia fisiológica es que la afinidad por el  $\text{NAD}^+$  *in vivo*, medida como la concentración del nucleótido necesaria para que la enzima esté al 50% formando el complejo ternario productivo, dependerá de la concentración intracelular del aldehído, de manera que la concentración fisiológica de 0.6 mM puede ser o no saturante para la enzima. Lo será sólo si el aldehído está presente a bajas concentraciones relativas a su  $K_m$ , lo que es de esperarse, y por tanto puede también variar con el aldehído que la enzima esté oxidando. Es por esto que, para poder inferir la posible actividad de la enzima *in vivo*, es necesario determinar el mecanismo cinético de la enzima con el sustrato fisiológico, una vez que lo conozcamos.

## VIII CONCLUSIONES

- La enzima *PA2125* mostró preferencia por aldehídos aromáticos y una actividad nula o despreciable frente a aldehídos no aromáticos, incluidos el BAL y el furfural. Debido a que el contexto genómico del gen *pa2125* es diferente al de los genes que codifican para otras enzimas ALDHs involucradas en la degradación de aldehídos aromáticos, *PA2125* podría participar en una ruta metabólica de degradación aún no conocida de compuestos aromáticos contaminantes de suelos para los que ya se conocen rutas de degradación, o de algún compuesto aromático contaminante cuya degradación aún no se conoce. Nuestros resultados no nos permiten concluir con certeza cuál podría ser ese compuesto, aunque sugieren que podría ser benzaldehído di o tri sustituido por grupos hidroxilos.
- La enzima *PA2125* pertenece a una familia dentro de la superfamilia de las ALDHs que no había sido estudiada y cuyos miembros, enzimas todas de proteobacterias que pueden crecer en suelo, poseen funciones fisiológicas que han sido en su mayoría mal anotadas o no conocidas. Probablemente todas las enzimas de esta familia, que pensamos se agrupan en al menos cuatro subfamilias, tengan como sustratos aldehídos aromáticos.
- De acuerdo con nuestros resultados de análisis de secuencias, modelado por homología y acoplamiento molecular, existen cuatro residuos en el sitio de unión del aldehído de la *PA2125* que juegan un papel importante en determinar la especificidad por este sustrato: Ser150, Arg157, Phe281 y Glu449. Estos residuos son raros en otras enzimas ALDHs y pueden considerarse como firmas de esta familia y de las subfamilias que la forman.

- El residuo Gly175 en el sitio de unión del nucleótido permite la unión del grupo 2'-fosfato de la ribosa de la adenosina del NADP<sup>+</sup> sin presentar impedimentos estéricos o repulsiones electrostáticas. Sin embargo, el acoplamiento molecular de los dos nucleótidos al sitio activo del modelo por homología no nos permitió explicar el por qué el NAD<sup>+</sup> es mejor sustrato que el NADP<sup>+</sup>.
- El mecanismo cinético de la enzima PA2125 presenta características de posible relevancia para la especificidad por el aldehído, ya que la afinidad de la enzima por éste y el cofactor dependen no solo del tipo de aldehído sino también de las concentraciones fisiológicas de ambos.

## IX PERSPECTIVAS

- Determinar los parámetros cinéticos aparentes a concentraciones fisiológicas de  $\text{NAD}^+$  de otros posible sustratos de la *PA2125*, tales como gentisaldehído (2,5-dihidroxibenzaldehído), 3,4-dihidroxibenzaldehído, u algún otro que pueda surgir de una búsqueda más exhaustiva de compuestos aromáticos contaminantes, para compararlos con los probados en esta tesis y conocer si pueden o no ser sustratos.
- Determinar el mecanismo cinético de las reacciones  $\text{NAD}^+$ - y  $\text{NADP}^+$ -dependientes catalizadas por la *PA2125* con el mejor sustrato que encontremos, determinando los patrones de velocidad inicial, de inhibición por producto y de inhibición por análogos de sustratos.
- Realizar experimentos de cinética rápida con el aldehído de los anteriormente referidos que resulte ser el mejor sustrato de *PA2125*, y así completar su caracterización bioquímica.
- Obtener la estructura cristalizada de la enzima *PA2125* en su forma libre y en complejo tanto con  $\text{NAD(P)}^+$  como con aldehídos, con el fin de conocer bien las características del sitio activo, incluido el sitio de unión al cofactor, que determinan la especificidad de esta enzima por el cofactor y los aldehídos.
- Hacer mutagénesis sitio dirigida y la caracterización cinética de las enzimas mutantes que se obtengan con el fin de comprobar el papel que juegan los residuos Ser150, Arg157, Phe281 y Glu449, o cualquier otro residuo del sitio activo que sugiera la estructura cristalográfica, en la especificidad por el sustrato aldehído.
- Investigar la participación de la enzima *PA2125* en la degradación de diferentes compuestos aromáticos (benceno, tolueno, naftaleno, alcohol bencílico mono o di



sustituido por grupos -OH) comprobando si una mutante nula inhibe o anula el crecimiento de la bacteria en medios de cultivo que contengan como única fuente de carbono los compuestos en cuya degradación queremos investigar si esta enzima participa, en base en los resultados obtenidos en este trabajo.

- Realizar un estudio filogenético de la familia ALDH a la que pertenece la PA2125 con el fin de conocer las relaciones evolutivas de sus miembros y si ha existido presión de selección para determinar la especificidad por el aldehído.

## X BIBLIOGRAFÍA

- Aloush, V., Navon-Venezia, S., Seigmen-Igra, Y., Cabili, S. & Carmeli, Y. (2006). Multidrug resistant *Pseudomonas aeruginosa*: risk factors and clinical impact. *Antimicrob. Agents Chemother.* 50(1): 43-8.
- Armour, A. D., Shankowsky, H. A., Swanson, T., Lee, J. & Tredget, E. E. (2007). The impact of nosocomially-acquired resistant *Pseudomonas aeruginosa* infection in a burn unit. *J. Trauma.* 63(1): 164-71.
- Ashibe, B., Hirai, T., Higashi, K., Sekimizu, K. & Motojima, K. (2007). Dual subcellular localization in the endoplasmic reticulum and peroxisomes and a vital role in protecting against oxidative stress of fatty aldehyde dehydrogenase are achieved by alternative splicing. *J. Biol. Chem.* 282(28): 20763–20773.
- Aziz, R. K., Bartels, D., Best, A. A., DeJongh, M., Disz, T., Edwards, R. A., Formsma, K., Gerdes, S., Glass, E. M., Kubal, M., Meyer, F., Olsen, G. J., Olson, R., Osterman, A. L., Overbeek, R. A., McNeil, L. K., Paarmann, D., Paczian, T., Parrello, B., Pusch, G. D., Reich, C., Stevens, R., Vassieva, O., Vonstein, V., Wilke, A. & Zagnitko, O. (2008). The RAST Server: Rapid annotations using subsystems technology. *BMC Genomics* 9: 75.
- Barnsley, E. A. (1976). Role and regulation of the *ortho* and *meta* pathways of catechol metabolism in *pseudomonads* metabolizing naphthalene and salicylate. *J. Bacteriol.* 125(2): 404-408.
- Bergmans, D., Bonten, M., van Tiel, F., Gaillard, C., van der Geest, S., Wilting, R., de Leeuw, P. W. & Stobberingh, E. (1998). Cross-colonisation with *Pseudomonas aeruginosa* of patients in an intensive care unit. *Thorax* 53(12): 1053–1058.
- Bianconi, I., Jeukens, J., Freschi, L., Alcalá-Franco, B., Facchini, M., Boyle, B., Molinaro, A., Kukavica-Ibrulj, I., Tümmler, B., Levesque, R. C. & Bragonz, A. (2015). Comparative genomics and biological characterization of sequential *Pseudomonas aeruginosa* isolates from persistent airways infection. *BMC Genomics*, 16: 1105.
- Bicknell, L. S., Pitt, J., Aftimos, S., Ramadas, R., Maw, M. A. & Robertson, S. P. (2008). A missense mutation in ALDH18A1, encoding delta1-pyrroline-5-carboxylate synthase (P5CS), causes an autosomal recessive neurocutaneous syndrome. *Eur. J. Hum. Genet.* 16: 1176–1186.

- Black, W. & Vasiliou, V. (2009). The aldehyde dehydrogenase gene superfamily resource center. *Hum. Genomics* 4(2): 136–142.
- Boch, J., Nau-Wagner, G., Kneip, S., Bremer, E. (1997). Glycine betaine aldehyde dehydrogenase from *Bacillus subtilis*: characterization of an enzyme required for the synthesis of the osmoprotectant glycine betaine. *Arch. Microbiol.* 168: 282–289.
- Bodey, G. P. (2001). *Pseudomonas aeruginosa* infections in cancer patients: have they gone away? *Curr. Opin. Infect. Dis.* 14(4): 403-7.
- Bradford, M. M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* 72: 248-254.
- Bredenstein, E. B. M., de la Fuente-Nunez, C. & Hancock, R. E. W. (2011). *Pseudomonas aeruginosa*: all roads lead to resistance. *Trends Microbiol.* 19(8): 419-25.
- Brocker, C., Lassen, N., Estey, T., Pappa, A., Cantore, M., Orlova, V. V., Chavakis, T., Kavanagh, K.L., Oppermann, U. & Vasiliou, V. (2010). Aldehyde dehydrogenase 7A1 (ALDH7A1) is a novel enzyme involved in cellular defense against hyperosmotic stress. *J. Biol. Chem.* 285: 18452–18463.
- Brown, M. E. & Chang, M. C. (2014). Exploring bacterial lignin degradation. *Curr. Opin. Biotechnol.* 19: 1–7.
- Brunner, N. A., Brinkmann, H., Siebers, B. & Hensel, R. (1998). NAD<sup>+</sup>-dependent glyceraldehyde-3-phosphate dehydrogenase from *Thermoproteus tenax* the first identified archaeal member of the aldehyde dehydrogenase superfamily is a glycolytic enzyme with unusual regulatory properties. *J. Biol. Chem.* 273: 6149–6156.
- Bugg, T. D., Ahmad, M., Hardiman, E. M. & Singh, R. (2011). The emerging role for bacteria in lignin degradation and bio-product formation. *Curr. Opin. Biotechnol.* 22, 394–400.
- Bugg, T. D., Ahmad, M., Hardiman, E. M. & Rahmanpour, R. (2011). Pathways for degradation of lignin in bacteria and fungi. *Nat. Prod. Rep.* 28: 1883–1896.
- Burley, S. K., Berman, H. M., Christie, C., Duarte, J., Feng, Z., Westbrook, J., Young, J. & Zardecki, C. (2017). RCSB Protein Data Bank: Sustaining a living digital data

- resource that enables breakthroughs in scientific research and biomedical education. *Protein Sci.* doi: 10.1002/pro.3331.
- Butler, C. S. & Mason, J. R. (1997). Structure–function analysis of the bacterial aromatic ring-hydroxylating dioxygenases. *Adv. Microb. Physiol.* 38: 47–84.
- Calderón, C. E., Ramos, C., de Vicente, A. & Cazorla, F. M. (2015). Comparative genomic analysis of *Pseudomonas chlororaphis* PCL1606 reveals new insight into antifungal compounds involved in biocontrol. *Mol. Plant Microbe. Interact.* 28(3): 249-260.
- Capestany, G. J., McDaniels, J. & Opgrande, J. I. (1977). The effect of sulphate on biological treatment of phenol-benzaldehyde wastes. *J. Water Pollut. Control Fed.* 49: 256-261.
- Cardona-Cardona, Y. V. (2017). Clonación, sobreexpresión y caracterización cinética y estructural de la enzima aminoaldehído deshidrogenasa PA5312 (PaKauB) de *Pseudomonas aeruginosa* (Tesis para optar el grado de maestría). Universidad Nacional Autónoma de México, Ciudad de México, México.
- Caspi, R., Billington, R., Fulcher, C. A., Keseler, I. M., Kothari, A., Krummenacker, M., Latendresse, M., Midford, P. E., Ong, Q., Ong, W. K., Paley, S., Subhraveti, P. & Karp, P. D. (2017). The MetaCyc database of metabolic pathways and enzymes. *Nucleic Acids Res.* doi: 10.1093/nar/gkx935.
- Cerniglia, C. E. (1992). Biodegradation of polycyclic aromatic hydrocarbons. *Biodegradation* 3: 351–368.
- Chen, J., Qin, J., Zhu, Y-G., de Lorenzo, V. & Rosen, B. P. (2013). Engineering the soil bacterium *Pseudomonas putida* for arsenic methylation. *Appl. Environ. Microbiol.* 79(14): 4493–4495.
- Cho, J-C., & Tiedje, J. M. (1999) Biogeography and Degree of Endemicity of Fluorescent *Pseudomonas* Strains in Soil. *Appl. Environ. Microbiol.* 66: 5448-5456.
- Chou, H. T., Li, J. Y., Peng, Y. C. Lu, C. D. (2013). Molecular characterization of PauR and its role in control of putrescine and cadaverine catabolism through the  $\gamma$ -glutamylolation pathway in *Pseudomonas aeruginosa* PAO1. *J. Bacteriol.* 195: 3906–3913.

- Chowdhury, P. P., Sarkar, J., Basu, S. & Dutta, T. K. (2014). Metabolism of 2-hydroxy-1-naphthoic acid and naphthalene via gentisic acid by distinctly different sets of enzymes in *Burkholderia sp.* strain BC1. *Microbiology* 160(Pt 5): 892-902.
- Cleland, W. W. (1963). The kinetics of enzyme-catalyzed reactions with two or more substrates or products. *Biochim. Biophys. Acta* 67: 104–137.
- Coitinho, J. B.; Pereira, M. S.; Costa, D. M.; Guimaraes, S. L.; Araujo, S. S.; Hengge, A. C.; Brandao, T. A.; Nagem, R. A. P. (2016). Structural and kinetic properties of the aldehyde dehydrogenase NahF, a broad substrate specificity enzyme for aldehyde oxidation. *Biochemistry* 55: 5453-5463.
- Cobessi, D., Tête-Favier, F., Marchal, S., Azza, S., Branlant, G. & Aubry, A. (1999). Apo and holo crystal structures of an NADP-dependent aldehyde dehydrogenase from *Streptococcus mutans*. *J. Mol. Biol.* 290: 161-73.
- D'Ambrosio, K., Pailot, A., Talfournier, F., Didierjean, C., Benedetti, E., Aubry, A., Branlant, G. & Corbier, C.
- Dagley, S., Evans, W. C. & Ribbons, D. W. (1960). New pathways in the oxidative metabolism of aromatic compounds by microorganisms. *Nature* 188: 560–566.
- Díaz-Sánchez, A. G., González-Segura, L., Rudiño-Piñera, E., Lira-Rocha, A., Torres-Larios, A. & Muñoz-Clares, R. A. (2011). Novel NADPH-cysteine covalent adduct found in the active site of an aldehyde dehydrogenase. *Biochem. J.* 439(3): 443-52.
- Díaz-Sánchez, A. G., González-Segura, L., Mújica-Jiménez, C., Rudiño-Piñera, E., Montiel, C., Martínez-Castilla, L. P. & Muñoz-Clares, R. A. (2012). Amino acid residues critical for the specificity for betaine aldehyde of the plant ALDH10 isoenzyme involved in the synthesis of glycine betaine. *Plant Physiol.* 158: 1570-1582.
- Dowling, D. N. & O’Gara, F. (1994). Metabolites of *Pseudomonas* involved in the biocontrol of plant disease. *Trends Biotechnol.* 12(4): 133-141.
- Dubendorff, J. W. & Studier, F. W. (1991). Controlling basal expression in an inducible T7 expression system by blocking the target T7 promoter with lac repressor. *J. Mol. Biol.* 219: 45-59.

- Elomari, M., Coroler, L., Hoste, B., Gillis, M. & Izard, D. (1996). DNA relatedness among *Pseudomonas* strains isolated from natural mineral waters and proposal of *Pseudomonas veronii* sp. nov. *Int. J. Syst. Bacteriol.* 46: 1138–1144.
- Emsley, P., & Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Cryst. D.* 60: 2126-2132.
- Estahbanati, H. K., Kashani, P. P. & Ghanaatpisheh, F. (2002). Frequency of *Pseudomonas aeruginosa* serotypes in burn wound infections and their resistance to antibiotics. *Burns* 28(4): 637-41.
- Eswar, N., Eramian, D., Webb, B., Shen, M. & Sali, A. (2008). Protein structure modeling with MODELLER. *Methods Mol. Biol.* 426: 145-59.
- Falkenberg, P. & Strøm A.R. (1990). Purification and characterization of osmoregulatory betaine aldehyde dehydrogenase of *Escherichia coli*. *Biochim. Biophys. Acta* 1034: 253-259.
- Feldman, R. I., y H. Weiner. (1972). Horse liver aldehyde dehydrogenase. II. Kinetics and mechanistic implications of the dehydrogenase and esterase activity. *J. Biol. Chem.* 247: 267-272.
- Ferraris, G. & Courerot, L. (2004). Evaluación de la utilización de bacterias solubilizadoras de fósforo (*Pseudomonas spp*) en trigo. *Rev. Tecnol. Agropecuaria* 9: 37-39.
- Figuroa-Soto, C. G., & Valenzuela-Soto, E. M. (2000). Kinetic study of porcine kidney betaine aldehyde dehydrogenase. *Biochem. Biophys. Res. Commun.* 269: 596–603.
- Fujitani, S., Sun, H. Y., Yu, V. L. & Weingarten, J. A. (2011). Pneumonia due to *Pseudomonas aeruginosa*: part I: epidemiology, clinical diagnosis, and source. *Chest* 139(4): 909-919.
- Garza-Ramos, G., Mújica-Jiménez, C. & Muñoz-Clares, R. A. (2013). Potassium and ionic strength effects on the conformational and thermal stability of two aldehyde dehydrogenases reveal structural and functional roles of K<sup>+</sup>-binding sites. *PLoS One.* 8(1): e54899.
- Gasteiger, E. E., Gattiker, A., Hoogland, C., Ivanyi, I., Appel, R. D. & Bairoch, A. (2003). ExPASy: The proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res.* 31(13): 3784-8.

- Gill, S. C. & Von Hippel, P. H. (1989). Calculation of protein extinction coefficients from amino acid sequence data. *Anal. Biochem.* 182: 319-326.
- González-Segura, L., Velasco-García, R. & Muñoz-Clares, R. A. (2002). Modulation of the reactivity of the essential cysteine residue of betaine aldehyde dehydrogenase from *Pseudomonas aeruginosa*. *Biochem. J.* 361: 577-585.
- González-Segura, L., Rudiño-Piñera, E., Muñoz-Clares, R. A. & Horjales, E. (2009). The crystal structure of a ternary complex of betaine aldehyde dehydrogenase from *Pseudomonas aeruginosa* provides new insight into the reaction mechanism and shows a novel binding mode of the 2'-phosphate of NADP<sup>+</sup> and a novel cation binding site. *J. Mol. Biol.* 385: 542-557.
- González-Segura, L., Riveros-Rosas, H., Julián-Sánchez, A. & Muñoz-Clares, R. A. (2015). Residues that influence coenzyme preference in the aldehyde dehydrogenases. *Chem.-Biol. Interact.* 234: 59–74.
- Haas, D. & Keel, C. (2003). Regulation of antibiotic production in root-colonizing *Pseudomonas* spp. and relevance for biological control of plant disease. *Annu. Rev. Phytopathol.* 41: 117-153.
- Haigler, B. E. & Spain, J. C. (1993). Biodegradation of 4-nitrotoluene by *Pseudomonas* sp. strain 4NT. *Appl. Environ. Microbiol.* 59(7): 2239-2243.
- 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.
- Hanukoglu, I. (2015). Proteopedia: Rossmann fold: A beta-alpha-beta fold at dinucleotide binding sites. *Biochem. Mol. Biol. Educ.* 43(3): 206-209.
- Harayama, S., Kok, M. & Neidle, E. L. (1992). Functional and evolutionary relationships among diverse oxygenases. *Ann. Rev. Microbiol.* 46: 565–601.
- Hayaishi, O. (2008). From oxygenase to sleep. *J. Biol. Chem.* 283: 19165–19175.
- Hempel, J., Nicholas, H. & Lindahl, R. (1993). Aldehyde dehydrogenases: Widespread structural and functional diversity within a shared framework. *Prot. Sci.* 2: 1890-1900.
- Henehan, G. T. & Tipton, K. F. (1992). Steady-state kinetic analysis of aldehyde dehydrogenase from human erythrocytes. *Biochem. J.* 287: 145-150.

- Humphrey, W., Dalke, A., & Schulten, K. (1996). VMD: visual molecular dynamics. *J. Mol. Graph.* 14(1): 33-38.
- Inoue, J., Shaw, J. P., Reikik, M. & Harayama, S. (1995). Overlapping substrate specificities of benzaldehyde dehydrogenase (the xylC gene product) and 2-hydroxymuconic semialdehyde dehydrogenase (the xylG gene product) encoded by TOL plasmid pWW0 of *Pseudomonas putida*. *J. Bacteriol.* 177(5): 1196-1201.
- Itoh, Y. (1997). Cloning and characterization of the *aru* genes encoding enzymes of the catabolic arginine succinyltransferase pathway in *Pseudomonas aeruginosa*. *J. Bacteriol.* 179(23): 7280-90.
- James, K. D. & Williams, P. A. (1998). ntn genes determining the early steps in the divergent catabolism of 4-nitrotoluene and toluene in *Pseudomonas* sp. strain TW3. *J. Bacteriol.* 180(8): 2043-2049.
- James, K. D., Hughes, M. A. & Williams, P. A. (2000). Cloning and expression of ntnD, encoding a novel NAD(P)<sup>+</sup>-independent 4-nitrobenzyl alcohol dehydrogenase from *Pseudomonas* sp. Strain TW3. *J. Bacteriol.* 182(11): 3136-41.
- Jann, A., Stalon, V., Wauven, C. V., Leisinger, T. & Haas, D. (1986). N-Succinylated intermediates in an arginine catabolic pathway of *Pseudomonas aeruginosa*. *Proc. Natl. Acad. Sci. U.S.A.* (13): 4937-41.
- Jakoby, W. J. (1963). Aldehyde dehydrogenases. Boyer, P. D. (ed). *The enzymes* New York, Academic Press. 7, 2nd ed. pp. 203-221.
- Johansson, K., El-Ahmad, M., Ramaswamy, S., Hjelmqvist, L., Jörnvall, H. & Eklund, H. (1998). Structure of betaine aldehyde dehydrogenase at 2.1 Å resolution. *Protein Sci.* 87: 2106-2117.
- Krupenko, N. I., Holmes, R. S., Tsybovsky, Y. & Krupenko, S. A. (2015). Aldehyde dehydrogenase homologous folate enzymes: evolutionary switch between cytoplasmic and mitochondrial localization. *Chem.-Biol. Interact.* 234: 12–17.
- Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head bacteriophage T4. *Nature* 227: 680-685.
- Lambert, P. A. (2002). Mechanisms of antibiotic resistance in *Pseudomonas aeruginosa*. *J. R. Soc. Med.* 95(suppl 41): 22-6.



- Lamark, T., Røkenes, T. P., McDougall, J. & Strøm, A. R. (1997). The complex bet promoters of *Escherichia coli*: regulation by oxygen (ArcA), choline (BetI), and osmotic stress. *J. Bacteriol.* 178: 1655-1662.
- Lambert, P. A. (2002). Mechanisms of antibiotic resistance in *Pseudomonas aeruginosa*. *J. R. Soc. Med.* 95(suppl 41): 22-6.
- Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., Thompson, J. D., Gibson, T. J. & Higgins, D. G. (2007). Clustal W and Clustal X version 2.0. *Bioinformatics* 23: 2947-2948.
- Lindahl, R. (1992). Aldehyde dehydrogenases and their role in carcinogenesis. *Crit. Rev. Biochem. Mol.* 27: 283-335.
- Liu, Z. J., Sun, Y. J., Rose, J., Chung, Y. J., Hsiao, C. D., Chang, W. R., Kuo, I., Perozich, J., Lindahl, R. & Hempel, J. (1997). The first structure of an aldehyde dehydrogenase reveals novel interactions between NAD<sup>+</sup> and the Rossmann fold. *Nat. Struct. Biol.* 4: 317–326.
- Liu, F., & Schnable, P. S. (2002). Functional specialization of maize mitochondrial aldehyde dehydrogenases. *Plant Physiol.* 130(4): 1657–1674.
- MacGibbon, A. K., Buckey, P. D. & Blackwell, L. F. (1977). Evidence for two-step binding of reduced nicotinamide-adenine dinucleotide to aldehyde dehydrogenase. *Biochem. J.* 165: 455–462.
- Marchler-Bauer, A., Lu, S., Anderson, J. B., Chitsaz, F., Derbyshire, M. K., DeWeese-Scott, C., Fong, J. H., Geer, L. Y., Geer, R. C., Gonzales, N. R., Gwadz, M., Hurwitz, D. I., Jackson, J. D., Ke, Z., Lanczycki, C. J. & Lu, F. (2011). CDD: a Conserved Domain Database for the functional annotation of proteins. *Nucleic Acids Res.* 39(Database issue): D225–D229.
- Marquardt, D. W. (1963). An algorithm for least-squares estimation of nonlinear parameters. *J. Soc. Ind. Appl. Math.* 11: 431-441.
- Martínez-Castilla, L. P., & Rodríguez-Sotres, R. (2010). A score of the ability of a threedimensional protein model to retrieve its own sequence as a quantitative measure of its quality and appropriateness. *PLoS ONE* 5(9): e12483.

- Masai, E., Katayama, Y. & Fukuda, M. (2007). Genetic and biochemical investigations on bacterial catabolic pathways for lignin-derived aromatic compounds. *Biosci. Biotechnol. Biochem.* 71: 1–15.
- McBride, L. J. & Caruthers, M. H. (1983). An investigation of several deoxynucleoside phosphoramidites useful for synthesizing deoxyoligonucleotides. *Tetrahedron Lett.* 24: 245-248.
- Mittal, R., Aggarwal, S., Sharma, S., Chhibber, S. & Harjai, K. (2009). Urinary tract infections caused by *Pseudomonas aeruginosa*: a mini review. *J. Infect. Public Health* 2(3): 101-11.
- Moore, S. M., Liang, T., Graves, T. J., McCall, K. M., Carr, L. G. & Ehlers, C. L. (2009). Identification of a novel cytosolic aldehyde dehydrogenase allele, ALDHIA1\*4. *Hum. Genomics* 3(4): 304–307.
- Mori, N., Kawakami, B., Hyakutome, K., Tani, Y. & Yamada, H. (1980). Characterization of betaine aldehyde dehydrogenase from *Cylindrocarpon didymium* M-1. *Agric. Biol. Chem.* 40: 3015-3016.
- Morris, C. E., Sands, D. C., Vanneste, J. L., Montarry, J., Oakley, B., Guilbaud, C. & Glaux, C. (2010). Inferring the evolutionary history of the plant pathogen *Pseudomonas syringae* from its biogeography in headwaters of rivers in North America, Europe, and New Zealand. *mBio* 1: e00107-10.
- Morris, N. S., Stickler, D. J. & McLean, R. J. (1999). The development of bacterial biofilms on indwelling urethral catheters. *World J. Urol.* 17(6): 345-50.
- Mullis K, Faloona F, Scharf S, Saiki R, Horn G, Erlich H. (1986). Specific enzymatic amplification of DNA in vitro: The polymerase chain reaction. *Cold Spring Harb. Symp. Quant. Biol.* 51 (Pt 1): 263-273.
- Murray, J. L., Kwon, T., Marcotte, E. M. & Whiteley, M. (2015). Intrinsic antimicrobial resistance determinants in the superbug *Pseudomonas aeruginosa*. *MBio* 6(6): e01603-15.
- Muñoz-Clares, R. A., Díaz-Sánchez, A. G., González-Segura, L., & Montiel, C. (2010). Kinetic and structural features of betaine aldehyde dehydrogenases: mechanistic and regulatory implications. *Arch. Biochem. Biophys.* 493: 71-81.

- Muñoz-Clares, R. A., González-Segura, L. & Díaz-Sánchez, Á. G. (2011). Crystallographic evidence for active-site dynamics in the hydrolytic aldehyde dehydrogenases. Implications for the deacylation step of the catalyzed reaction. *Chem.-Biol. Interact.* 191: 137-146.
- Muñoz-Clares, R. A., González-Segura, L., Murillo-Melo, D. S. & Riveros-Rosas, H. (2017). Mechanisms of protection against irreversible oxidation of the catalytic cysteine of ALDH enzymes: possible role of vicinal cysteines. *Chem-Biol. Interact.* 276: 52-64.
- Nagasawa, T., Kawabata, Y., Tani, Y. & Ogata, K. (1976). Purification and characterization of betaine aldehyde dehydrogenase from *Pseudomonas aeruginosa* A-16. *Agric. Biol. Chem.* 40: 1743-1749.
- Nagradova, N. K., (2001). Study of the properties of phosphorylating D-glyceraldehyde-3-phosphate dehydrogenase. *Biochemistry (Mosc)*. 66(10): 1067-76.
- Nichols, N. N., & Mertens, J. A. (2008). Identification and transcriptional profiling of *Pseudomonas putida* genes involved in furoic acid metabolism. *FEMS Microbiol. Lett*, 284(1): 52-57.
- Nixon, G. M., Armstrong, D. S., Carzino, R., Carlin, J. B., Olinsky, A., Robertson, C. F. & Grimwood, K. (2001). Clinical outcome after early *Pseudomonas aeruginosa* infection in cystic fibrosis. *J. Pediatr.* 138(5): 699-704.
- Oberhardt, M. A., Puchalka, J., Fryer, K. E., Martins dos Santos, V. A. P. & Papin, J. A. (2008). Genome-scale metabolic network analysis of the opportunistic pathogen *Pseudomonas aeruginosa* PAO1. *J. Bacteriol.* 190: 2790-2803.
- Ornston, L. N. & Stanier, R. Y. (1996). The conversion of catechol and protocatechuate to  $\beta$ -keto adipate by *Pseudomonas putida*. *J. Biol. Chem.* 241: 3776–3786.
- Parker, C. M., Kutsogiannis, J., Muscedere, J., Cook, D., Dodek, P., Day, A. G. & Heyland, D. K. (2008). Ventilator-associated pneumonia caused by multidrug-resistant organisms or *Pseudomonas aeruginosa*: prevalence, incidence, risk factors, and outcomes. *J. Crit. Care* 23(1): 18-26.
- Pendleton, J. N., Gorman, S. P. & Gilmore, B. F. (2013). Clinical relevance of the ESKAPE pathogens. *Expert Rev. Anti Infect. Ther.* 11: 297–308.

- Pettersen, E. F., Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C. & Ferrin, T. E. 2004. UCSF Chimera: a visualization system for exploratory research and analysis. *J. Comput. Chem.* 25(13): 1605-12
- Raaijmakers, J. M., Vlami, M. & de Souza, J. T. (2002). Antibiotic production by bacterial biocontrol agents. *Antonie Van Leeuwenhoek* 81: 537-547.
- Rahme, L. G., Stevens, E. J., Wolfort, S. F., Shao, J., Tompkins, R. G. & Ausubel, F. M. (1995). Common virulence factor for bacterial pathogenicity in plants and animals. *Science* 268: 1899-1902.
- Raman, S., Vernon, R., Thompson, J., Tyka, M., Sadreyev, R., Pei, J., Kim, D., Kellogg, E., DiMaio, F., Lange, O., Kinch, L., Sheffler, W., Kim, B., Das, R., Grishin, N. V., & Baker, D. (2009). Structure prediction for CASP8 with all-atom refinement using Rosetta. *Proteins* 77(Suppl. 9): 89-99.
- Rathinasabapathi, B., McCue, K. F., Gage, D. A. & Hanson, A. D. (1996). Metabolic engineering of glycine betaine synthesis: plant betaine aldehyde dehydrogenases lacking typical transit peptides are targeted to tobacco chloroplasts where they confer betaine aldehyde resistance, *Planta* 193: 155–162.
- Rau, M. H., Marvig, R. L., Ehrlich, G. D., Molin, S. & Jelsbak, L. (2012). Deletion and acquisition of genomic content during early stage adaptation of *Pseudomonas aeruginosa* to a human host environment. *Environ. Microbiol.* 14(8): 2200-2211.
- Ressner, R. A., Murray, C. K., Griffith, M. E., Rasnake, M. S., Hospenthal, D. R. & Wolf, S. E. (2008). Outcomes of bacteremia in burn patients involved in combat operations overseas. *J. Am. Coll. Surg.* 206(3): 439-44.
- Revelles, O., Espinosa-Urgel, M., Fuhrer, T., Sauer, U. & Ramos, J. L. (2005). Multiple and interconnected pathways for L-lysine catabolism in *Pseudomonas putida* KT2440. *J. Bacteriol.* 7(21): 7500-10.
- Řezáč, J., Fanfrlík, J., Salahub, D., & Hobza, P. (2009). Semiempirical quantum chemical PM6 method augmented by dispersion and H-bonding correction terms reliably describes various types of noncovalent complexes. *J. Chem. Theory Comput.* 5(7): 1749–1760.
- Ritchie, C. D. & Sager, W. F. Progress in physical organic chemistry, Interscience publishers, a division of John Wiley Sons, Inc. United States of America, 1964.

- Riveros-Rosas, H., L. González-Segura, A. Julián-Sánchez, A. G. Díaz-Sánchez, y R. A. Muñoz-Clares. (2013). Structural determinants of substrate specificity in aldehyde dehydrogenases. *Chem-Biol. Interact.* 202: 51–61.
- Roberts, E., Eargle, J., Wright, D., & Luthey-Schulten, Z. (2006). MultiSeq: unifying sequence and structure data for evolutionary analysis. *BMC Bioinformatics* 7: 382
- Rolston, K. V. & Bodey, G. P. (1992). *Pseudomonas aeruginosa* infection in cancer patients. *Cancer Invest.* 10(1): 43-59.
- Rossmann, M. G., Moras, D., & Olsen, K. W. (1974). Chemical and biological evolution of nucleotide-binding protein. *Nature* 250: 194-199.
- Samuel, M. S., Sivaramakrishna, A. & Mehta, A. (2014). Bioremediation of p-Nitrophenol by *Pseudomonas putida* 1274 strain. *J. Environ. Health Sci. Eng.* 12: 53-60.
- Samuelson, J., Zhang, W. W., Kumar, A., Descoteaux, S., Shen, P. S. & Bailey, G. (1992). Primary structures of alcohol and aldehyde dehydrogenase genes of *Entamoeba histolytica*. *Arch. Med. Res.* 23: 31–33.
- Sanger, F., & Coulson, A. R. (1975). A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. *J. Mol. Biol.* 94: 441-448.
- Segel, I. H. Steady-state kinetics of multireactant enzymes, in: *Enzyme kinetics. Behavior and analysis of rapid equilibrium and steady-state enzyme systems*, John Wiley Sons, Inc. United States of America, IN, 1993.
- Sidhu, R. S. & Blair, A. H. (1975). Human liver aldehyde dehydrogenase. Kinetics of aldehyde oxidation. *J. Biol. Chem.* 250: 7899-7904.
- Shaw, J. P. & Harayama, S. (1990). Purification and characterisation of TOL plasmid-encoded benzyl alcohol dehydrogenase and benzaldehyde dehydrogenase of *Pseudomonas putida*. *Eur. J. Biochem.* 191(3): 705-714.
- Shepp, D. H., Tang, I. T., Ramundo, M. B. & Kaplan, M. K. (1994). Serious *Pseudomonas aeruginosa* infection in AIDS. *J. Acquir. Immune. Defic. Syndr.* 7(8): 823-31.
- Shimoni, E., Baasov, T., Ravid, U. & Shoham, Y. (2002). The trans-anethole degradation pathway in an *Arthrobacter sp.* *J. Biol. Chem.* 277(14): 11866-11872.
- Shone, C. C. & Fromm, H. J. (1981). Steady-state and pre-steady-state kinetics of coenzyme A linked aldehyde dehydrogenase from *Escherichia coli*. *Biochemistry* 20: 7494–7501.

- Singh, S., Brocker, C., Koppaka, V., Ying, C., Jackson, B., Matsumoto, A., Thompson, D. C. & Vasiliou, V. (2013). Aldehyde dehydrogenases in cellular responses to oxidative/electrophilic stress. *Free Radic. Biol. Med.* 56: 89–101.
- Smith, L. T., Pocard, J. A., Bernard, T., & Le Rudulier, D. (1988). Osmotic control of glycine betaine biosynthesis and degradation in *rhizobium meliloti*. *J. Bacteriol.* 170: 3142-3149.
- Smith, J. J., Travis, S. M., Greenberg, E. P. & Welsh, M. J. (1996). Cystic fibrosis airway epithelia fail to kill bacteria because of abnormal airway surface fluid. *Cell* 85(2): 229-36.
- Söding, J. (2005). Protein homology detection by HMM-HMM comparison. *Bioinformatic* 21(7): 951-60.
- Söding, J., Biegert, A., & Lupas, A. N. (2005). The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Res.* 33(Web Server issue): W244-8.
- Söding, J., Remmert, M., Biegert, A. & Lupas, A. N. (2006). HHsenser: exhaustive transitive profile search using HMM-HMM comparison. *Nucleic Acids Res.* 34(Web Server issue): W374-8.
- Sophos, N. A. & Vasiliou, V. (2003). Aldehyde dehydrogenase gene superfamily: the 2002 update. *Chem.-Biol. Interact.* 143-144: 5-22.
- Stagos, D., Chen, Y., Cantore, M., Jester, J. & Vasiliou, V. V. (2010). Corneal aldehyde dehydrogenases: multiple functions and novel nuclear localization. *Brain Res. Bull.* 81: 211–218.
- Trott, O., & Olson, A. J. (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.
- Valenzuela-Soto, E. M. & Muñoz-Clares, R. A. (1993). Betaine-aldehyde dehydrogenase from leaves of *Amaranthus hypochondriacus* L. exhibits an iso ordered bi bi steady state mechanism. *J. Biol. Chem.* 268: 23818-23823. Erratum: 1994, *Ibid*, 269: 4692.
- Vasiliou, V., Pappa, A., & Petersen, D. R. (2000). Role of aldehyde dehydrogenases in endogenous and xenobiotic metabolism. *Chem.-Biol. Interact.* 129(1–2): 1-19.

- Velasco-García, R., Mújica-Jiménez, C., Mendoza-Hernández, G. & Muñoz-Clares, R. A. (1999). Rapid purification and properties of betaine aldehyde dehydrogenase from *Pseudomonas aeruginosa*. *J. Bacteriol.* 181: 1292–1300.
- Velasco-García, R., González-Segura, L. & Muñoz-Clares, R. A. (2000). Steady-state kinetic mechanism of the NADP<sup>-</sup>- and NAD<sup>+</sup>-dependent reactions catalysed by betaine aldehyde dehydrogenase from *Pseudomonas aeruginosa*. *Biochem. J.* 352: 675-683.
- Velasco-García, R., Chacón-Aguilar, V. M., Hervert-Hernández, D. & Muñoz-Clares, R. A. (2003). Inactivation of betaine aldehyde dehydrogenase from *Pseudomonas aeruginosa* and *Amaranthus hypochondriacus* L. leaves by disulfiram. *Chem.-Biol. Interact.* 143–144: 149-158.
- Velasco-García R., Villalobos, M. A., Ramírez-Romero, M. A., Mújica-Jiménez, C., Iturriaga, G. & Muñoz-Clares, R. A. (2006). Betaine aldehyde dehydrogenase from *Pseudomonas aeruginosa*: cloning, over-expression in *Escherichia coli*, and regulation by choline and salt. *Arch. Microbiol.* 185: 14–22.
- Webb, B., & Sali, A. (2014). Comparative protein structure modeling using Modeller. *Curr. Protoc. Bioinformatics* 47: 5.6.1-32.
- Winsor, G. L., Griffiths, E. J., Lo, R., Dhillon, B. K., Shay, J. A. & Brinkman, F. S. (2016). Enhanced annotations and features for comparing thousands of *Pseudomonas* genomes in the *Pseudomonas* genome database. *Nucleic. Acids Res.* doi: 10.1093/nar/gkv1227.
- Woese, C. R. (1987). Bacterial Evolution. *Microbiol. Rev.* 51(2): 221-271.
- Worlitzch, D., Tarran, R., Ulrich, M., Schwab, U., Cekici, A., Meyer, K. C., Birrer, P., Bellon, G., Berger, J., Weiss, T., Botzenhart, K., Yankaskas, J. R., Randell, S., Boucher, R. C. & Döring, G. (2002). Effects of reduced mucus oxygen concentration in airway *Pseudomonas* infections of cystic fibrosis patients. *J. Clin. Invest.* 109(3): 317-25.
- Yao, X., He, W., Lu, C. D. (2011). Functional characterization of seven  $\gamma$ -glutamylpolyamine synthetase genes and the bauRABCD locus for polyamine and  $\beta$ -alanine utilization in *Pseudomonas aeruginosa* PAO1. *J. Bacteriol.* 93(15):3923-30.

- Yoshida, A., Rzhetsky, A., Hsu, L.C. & Chang, C. (1998). Human aldehyde dehydrogenase gene family. *Eur. J. Biochem.* 251: 549–557.
- Yoshida, Y., Sato, M., Kezuka, Y., Hasegawa, Y., Nagano, K., Takebe, J. & Yoshimura, F. (2016). Acyl-CoA reductase PGN\_0723 utilizes succinyl-CoA to generate succinate semialdehyde in a butyrate-producing pathway of *Porphyromonas gingivalis*. *Arch. Biochem. Biophys.* 596: 138-48.
- Zaldívar-Machorro, V. J., López-Ortiz, M., Demare, P., Regla, I. & Muñoz-Clares, R. A. (2011). The disulfiram metabolites S-methyl-N,N-diethyldithiocarbamoyl sulfoxide and S-methyl-N,N-diethylthiocarbamoyl sulfone irreversibly inactivate betaine aldehyde dehydrogenase from *Pseudomonas aeruginosa*, both in vitro and in situ, and arrest bacterial grow. *Biochimie* 93(2): 286-95.
- Zahniser, M. P.D., Shreenath, P., Kneen, M. M., Kreinbring, C. A., Petsko, G. A., Ringe, D., McLeish, M. J. (2017). Structure and mechanism of benzaldehyde dehydrogenase from *Pseudomonas putida* ATCC 12633, a member of the Class 3 aldehyde dehydrogenase superfamily. *Protein Eng. Des. Sel.* 30: 271-278.
- Zhao, D., McCaffery, P., Ivins, K.J., Neve, R. L., Hogan, P., Chin, W. W. & Dräger, U. C. (1996). Molecular identification of a major retinoic-acid-synthesizing enzyme, a retinaldehyde-specific dehydrogenase. *Eur. J. Biochem.* 240(1): 15-22.



## **XI ANEXOS**

**Anexo 1. Genes que codifican para enzimas de la familia ALDH cd07120 contenidos en la base de datos del NCBI.** Los datos de esta tabla se obtuvieron por medio de BLASTP (<https://blast.ncbi.nlm.nih.gov/blast.cgi>).

Código de acceso	Descripción (Organismo)	Identidad (%)	Similitud (%)	Longitud alineamiento
SCY80161.1	Betaína Aldehído deshidrogenasa [ <i>Acinetobacter baumannii</i> ]	99	100	482
KTK92484.1	Aldehído deshidrogenasa [ <i>Pseudomonas putida</i> ]	99	99	482
SAJ30947.1	putative Aldehído deshidrogenasa [ <i>Enterobacter cloacae</i> ]	99	99	482
WP_011356651.1	Aldehído deshidrogenasa [ <i>Burkholderia lata</i> ]	79	89	476
WP_053260877.1	Aldehído deshidrogenasa [ <i>Pseudomonas chlororaphis</i> ]	79	89	481
WP_025405893.1	Aldehído deshidrogenasa [ <i>Burkholderia thailandensis</i> ]	78	88	479
WP_060360071.1	Aldehído deshidrogenasa [ <i>Burkholderia territorii</i> ]	77	87	477
WP_060009927.1	Aldehído deshidrogenasa [ <i>Burkholderia ubonensis</i> ]	77	87	479
WP_060324382.1	Aldehído deshidrogenasa [ <i>Burkholderia diffusa</i> ]	77	87	479
WP_059923624.1	Aldehído deshidrogenasa [ <i>Burkholderia stagnalis</i> ]	77	87	478
WP_059695047.1	Aldehído deshidrogenasa [ <i>Burkholderia cepacia</i> ]	77	87	481
WP_089342201.1	Aldehído deshidrogenasa [ <i>Burkholderia singularis</i> ]	76	86	478
WP_053526448.1	Aldehído deshidrogenasa [ <i>Burkholderia cenocepacia</i> ]	75	86	478
WP_094467823.1	Aldehído deshidrogenasa [ <i>Pseudomonas mandelii</i> ]	72	84	474
WP_011333787.1	Aldehído deshidrogenasa [ <i>Pseudomonas fluorescens</i> ]	71	82	479

WP_005787176.1	Aldehído deshidrogenasa [ <i>Pseudomonas synxantha</i> ]	71	83	473
EPF74764.1	Betaína aldehído deshidrogenasa [ <i>Acinetobacter gyllenbergii</i> CIP 110306]	69	86	227
WP_090290692.1	Aldehído deshidrogenasa [ <i>Pseudomonas brenneri</i> ]	69	83	476
WP_038447170.1	Aldehído deshidrogenasa [ <i>Pseudomonas fluorescens</i> ]	69	83	476
WP_076965093.1	Aldehído deshidrogenasa [ <i>Pseudomonas gessardii</i> ]	69	83	476
WP_032876309.1	Aldehído deshidrogenasa [ <i>Acinetobacter baumannii</i> ]	67	82	482
WP_070076747.1	Aldehído deshidrogenasa [ <i>Acinetobacter proteolyticus</i> ]	67	82	482
WP_049063726.1	Aldehído deshidrogenasa [ <i>Acinetobacter seifertii</i> ]	66	81	480
WP_013394782.1	Aldehído deshidrogenasa [ <i>Achromobacter xylosoxidans</i> ]	66	77	476
WP_081172951.1	Aldehído deshidrogenasa [ <i>Acinetobacter lactucae</i> ]	65	80	480
WP_061875655.1	Aldehído deshidrogenasa [ <i>Acinetobacter dijkschoorniae</i> ]	65	80	480
WP_088159312.1	Aldehído deshidrogenasa [ <i>Achromobacter xylosoxidans</i> ]	65	77	476
EKJ94815.1	Aldehído deshidrogenasa [ <i>Bradyrhizobium lupini</i> HPC(L)]	64	81	437
WP_073455189.1	Aldehído deshidrogenasa [ <i>Pseudonocardia thermophila</i> ]	64	75	477
WP_028733433.1	Aldehído deshidrogenasa [ <i>Rhizobium leguminosarum</i> ]	64	79	476
WP_052859957.1	Aldehído deshidrogenasa [ <i>Streptomyces celluloflavus</i> ]	63	79	480
WP_094792959.1	Aldehído deshidrogenasa [ <i>Streptomyces kasugaensis</i> ]	63	79	480

WP_040114747.1	Aldehído deshidrogenasa [ <i>Rhizobium gallicum</i> ]	63	78	479
WP_085994489.1	Aldehído deshidrogenasa [ <i>Rhizobium leguminosarum</i> ]	63	78	474
WP_074064792.1	Aldehído deshidrogenasa [ <i>Rhizobium etli</i> ]	63	78	477
WP_077979876.1	Aldehído deshidrogenasa [ <i>Rhizobium laguerreae</i> ]	63	78	479
WP_080830526.1	Aldehído deshidrogenasa [ <i>Agrobacterium genomosp.</i> 6]	63	78	477
WP_038614503.1	Aldehído deshidrogenasa [ <i>Rhizobium leguminosarum</i> ]	63	77	479
AHF85418.1	Aldehído deshidrogenasa [ <i>Rhizobium leguminosarum</i> bv. <i>trifolii</i> WSM1689]	63	77	479
WP_080865300.1	Aldehído deshidrogenasa [ <i>Agrobacterium tumefaciens</i> ]	63	78	476
KRG65434.1	Aldehído deshidrogenasa [ <i>Stenotrophomonas terrae</i> ]	62	77	480
WP_017373291.1	Aldehído deshidrogenasa [ <i>Enterobacteriaceae bacterium</i> LSJC7]	62	75	480
WP_058910710.1	Aldehído deshidrogenasa [ <i>Erwinia teleogrylli</i> ]	62	74	476
WP_017919246.1	Aldehído deshidrogenasa [ <i>Burkholderia gladioli</i> ]	61	74	475
WP_091616387.1	Aldehído deshidrogenasa [ <i>Amycolatopsis saalfeldensis</i> ]	61	75	477
WP_023172286.1	Aldehído deshidrogenasa [ <i>Gloeobacter kilauensis</i> ]	61	78	477
KEO59789.1	Aldehído deshidrogenasa [ <i>Thalassospira permensis</i> NBRC 106175]	61	77	476
WP_037987502.1	Aldehído deshidrogenasa [ <i>Thalassospira permensis</i> ]	61	77	476
WP_090024837.1	Aldehído deshidrogenasa [ <i>Chryseobacterium oleae</i> ]	61	77	475

WP_052143043.1	Aldehído deshidrogenasa [ <i>Burkholderia gladioli</i> ]	61	74	475
WP_024353461.1	Aldehído deshidrogenasa [ <i>Brevundimonas naejangsanensis</i> ]	60	78	478
WP_014986636.1	Aldehído deshidrogenasa [ <i>Nocardia brasiliensis</i> ]	60	74	474
APP18125.1	Gentisaldehído deshidrogenasa [ <i>Burkholderia sp. BC1</i> ]	60	75	477
WP_039140433.1	Aldehído deshidrogenasa [ <i>Flaviumibacter solisilvae</i> ]	60	77	474
WP_042255842.1	Aldehído deshidrogenasa [ <i>Nocardia brasiliensis</i> ]	60	74	474
WP_013078858.1	Aldehído deshidrogenasa [ <i>Caulobacter segnis</i> ]	60	74	476
WP_043677706.1	Aldehído deshidrogenasa [ <i>Nocardia vulneris</i> ]	59	73	474
WP_003891806.1	Aldehído deshidrogenasa [ <i>Mycobacterium smegmatis</i> ]	59	73	480
WP_090799963.1	Aldehído deshidrogenasa [ <i>Asanoa ishikariensis</i> ]	59	73	472
WP_048507183.1	Aldehído deshidrogenasa [ <i>Chryseobacterium angstadtii</i> ]	59	77	475
WP_093323662.1	Aldehído deshidrogenasa [ <i>Pseudosphingobacterium domesticum</i> ]	58	74	469
SEK90206.1	Betaína aldehído deshidrogenasa [ <i>Pseudosphingobacterium domesticum</i> ]	58	74	469
WP_047896639.1	Aldehído deshidrogenasa [ <i>Caballeronia mineralivorans</i> ]	58	72	472
WP_085119827.1	Aldehído deshidrogenasa [ <i>Serratia proteamaculans</i> ]	58	74	476
WP_074300019.1	Aldehído deshidrogenasa [ <i>Paraburkholderia phenazinium</i> ]	58	71	472
WP_012147298.1	Aldehído deshidrogenasa [ <i>Serratia proteamaculans</i> ]	58	74	476

WP_016336514.1	Aldehído deshidrogenasa [ <i>Amycolatopsis orientalis</i> ]	57	71	472
WP_042936682.1	Aldehído deshidrogenasa [ <i>Pseudomonas gingeri</i> ]	57	73	475
WP_073795725.1	Aldehído deshidrogenasa [ <i>Streptomyces uncialis</i> ]	57	72	472
WP_067824260.1	Aldehído deshidrogenasa [ <i>Nocardia inohanensis</i> ]	56	69	472
WP_047317165.1	Aldehído deshidrogenasa [ <i>Mycobacterium heraklionense</i> ]	56	73	471
WP_031043951.1	Aldehído deshidrogenasa [ <i>Streptomyces olivaceus</i> ]	56	70	473
WP_007497325.1	Aldehído deshidrogenasa [ <i>Streptomyces zinciresistens</i> ]	56	70	473
WP_074754022.1	Aldehído deshidrogenasa [ <i>Pseudomonas abietaniphila</i> ]	56	75	477
WP_021490172.1	Aldehído deshidrogenasa [ <i>Pseudomonas mendocina</i> ]	55	74	477
WP_030432593.1	Aldehído deshidrogenasa [ <i>Allokutzneria albata</i> ]	55	70	470
WP_059302984.1	Aldehído deshidrogenasa [ <i>Streptomyces canus</i> ]	55	70	473
WP_093783432.1	Aldehído deshidrogenasa [ <i>Streptomyces guanduensis</i> ]	55	71	472
WP_052840036.1	Aldehído deshidrogenasa [ <i>Kitasatospora aureofaciens</i> ]	55	70	473
WP_023536343.1	Aldehído deshidrogenasa [ <i>Streptomyces niveus</i> ]	55	69	473
WP_045312411.1	Aldehído deshidrogenasa [ <i>Lechevalieria aerocolonigenes</i> ]	55	71	472
WP_055470832.1	Aldehído deshidrogenasa [ <i>Streptomyces pathocidini</i> ]	55	68	472
WP_085018091.1	Aldehído deshidrogenasa [ <i>Cnuibacter physcomitrellae</i> ]	55	70	471

EFE78353.1	Aldehído deshidrogenasa [ <i>Streptomyces roseosporus</i> NRRL 15998]	55	69	473
WP_059008848.1	Aldehído deshidrogenasa [ <i>Streptomyces specialis</i> ]	54	67	472
WP_031940426.1	Aldehído deshidrogenasa [ <i>Rhodococcus defluvii</i> ]	54	69	471
WP_019633426.1	Aldehído deshidrogenasa [ <i>Actinomadura atramentaria</i> ]	54	68	475
WP_067682900.1	Aldehído deshidrogenasa [ <i>Nocardia miyunensis</i> ]	53	67	473
WP_052344701.1	Aldehído deshidrogenasa [ <i>Thermocrisum municipale</i> ]	53	67	479
WP_089105027.1	Aldehído deshidrogenasa [ <i>Streptomyces hyaluromycini</i> ]	53	69	479
WP_068567778.1	Aldehído deshidrogenasa [ <i>Tsukamurella pulmonis</i> ]	53	65	470
WP_036511165.1	Aldehído deshidrogenasa [ <i>Nocardia aobensis</i> ]	53	68	472
SDQ70049.1	Betaína aldehído deshidrogenasa [ <i>Thermomonospora chromogena</i> ]	53	67	471
WP_059210704.1	Aldehído deshidrogenasa [ <i>Streptomyces canus</i> ]	53	68	476
WP_015786915.1	Aldehído deshidrogenasa [ <i>Saccharomonospora viridis</i> ]	53	67	471
WP_040840209.1	Aldehído deshidrogenasa [ <i>Nocardia brevicatena</i> ]	53	68	473
WP_015795043.1	Aldehído deshidrogenasa [ <i>Catenulispora acidiphila</i> ]	53	68	477
WP_028931650.1	Aldehído deshidrogenasa [ <i>Pseudonocardia asaccharolytica</i> ]	52	66	471
WP_092555615.1	Aldehído deshidrogenasa [ <i>Actinoplanes derwentensis</i> ]	52	67	475
WP_071929043.1	Aldehído deshidrogenasa [ <i>Nocardia soli</i> ]	52	66	472

WP_063005819.1	Aldehído deshidrogenasa [ <i>Nocardia cummidelens</i> ]	51	68	472
WP_062974919.1	Aldehído deshidrogenasa [ <i>Nocardia flavorosea</i> ]	49	66	472
WP_066350715.1	Aldehído deshidrogenasa [ <i>Bordetella bronchialis</i> ]	49	62	481
WP_048436755.1	Aldehído deshidrogenasa [ <i>Methylobacterium platani</i> ]	49	62	480
WP_060847217.1	Aldehído deshidrogenasa [ <i>Methylobacterium aquaticum</i> ]	49	62	480
WP_055112563.1	Aldehído deshidrogenasa [ <i>Mycobacterium peregrinum</i> ]	49	65	472
WP_028213740.1	Aldehído deshidrogenasa [ <i>Paraburkholderia mimosarum</i> ]	49	62	479
WP_057928643.1	Aldehído deshidrogenasa [ <i>Burkholderia ambifaria</i> ]	49	62	479
WP_060091922.1	Aldehído deshidrogenasa [ <i>Burkholderia vietnamiensis</i> ]	49	62	479
WP_049286002.1	Aldehído deshidrogenasa [ <i>Ralstonia solanacearum</i> ]	49	62	479
WP_024888973.1	Aldehído deshidrogenasa [ <i>Luteimonas huabeiensis</i> ]	49	62	471
WP_061133346.1	Aldehído deshidrogenasa [ <i>Caballeronia fortuita</i> ]	48	62	479
WP_068574557.1	Aldehído deshidrogenasa [ <i>Tsukamurella pseudospumae</i> ]	48	65	469
WP_058833483.1	Aldehído deshidrogenasa [ <i>Luteimonas abyssii</i> ]	48	62	471
WP_018719436.1	Aldehído deshidrogenasa [ <i>Arhodomonas aquaeolei</i> ]	48	62	482
SDG68883.1	Acyl-CoA reductase [ <i>Streptomyces jietaisiensis</i> ]	48	67	474
WP_025359011.1	Aldehído deshidrogenasa [ <i>Kutzneria albida</i> ]	48	66	470



WP_009762628.1	Aldehído deshidrogenasa [ <i>Microvirga lotononidis</i> ]	48	66	468
WP_093973911.1	Aldehído deshidrogenasa [ <i>Boseongicola aestuarii</i> ]	48	61	482
WP_057589755.1	Aldehído deshidrogenasa [ <i>Variovorax paradoxus</i> ]	48	62	478
KOT37956.1	Aldehído deshidrogenasa [ <i>Streptomyces caelestis</i> ]	48	67	474
KEG43654.1	Aldehído deshidrogenasa [ <i>Streptomyces griseorubens</i> ]	48	67	474
WP_035529310.1	Aldehído deshidrogenasa [ <i>Paraburkholderia sacchari</i> ]	48	62	479
WP_085635662.1	Aldehído deshidrogenasa [ <i>Marivita geojedonensis</i> ]	48	62	471
WP_062963663.1	Aldehído deshidrogenasa [ <i>Nocardia africana</i> ]	48	65	469
WP_033277657.1	Aldehído deshidrogenasa [ <i>Actinospica acidiphila</i> ]	48	67	474
WP_018179912.1	Aldehído deshidrogenasa [ <i>Gordonia kroppenstedtii</i> ]	48	64	469
WP_041802429.1	Aldehído deshidrogenasa [ <i>Rhodopseudomonas palustris</i> ]	48	61	480
WP_033346353.1	Aldehído deshidrogenasa [ <i>Catenuloplanes japonicus</i> ]	47	65	470
WP_085632744.1	Aldehído deshidrogenasa [ <i>Marivita cryptomonadis</i> ]	47	61	471
WP_062361727.1	Aldehído deshidrogenasa [ <i>Variovorax paradoxus</i> ]	47	61	480
WP_043378184.1	Aldehído deshidrogenasa [ <i>Streptomyces luteus</i> ]	47	67	474
WP_032616514.1	Aldehído deshidrogenasa [ <i>Leclercia adecarboxylata</i> ]	47	66	454
WP_090929195.1	Aldehído deshidrogenasa [ <i>Palleronia marisminoris</i> ]	47	60	481

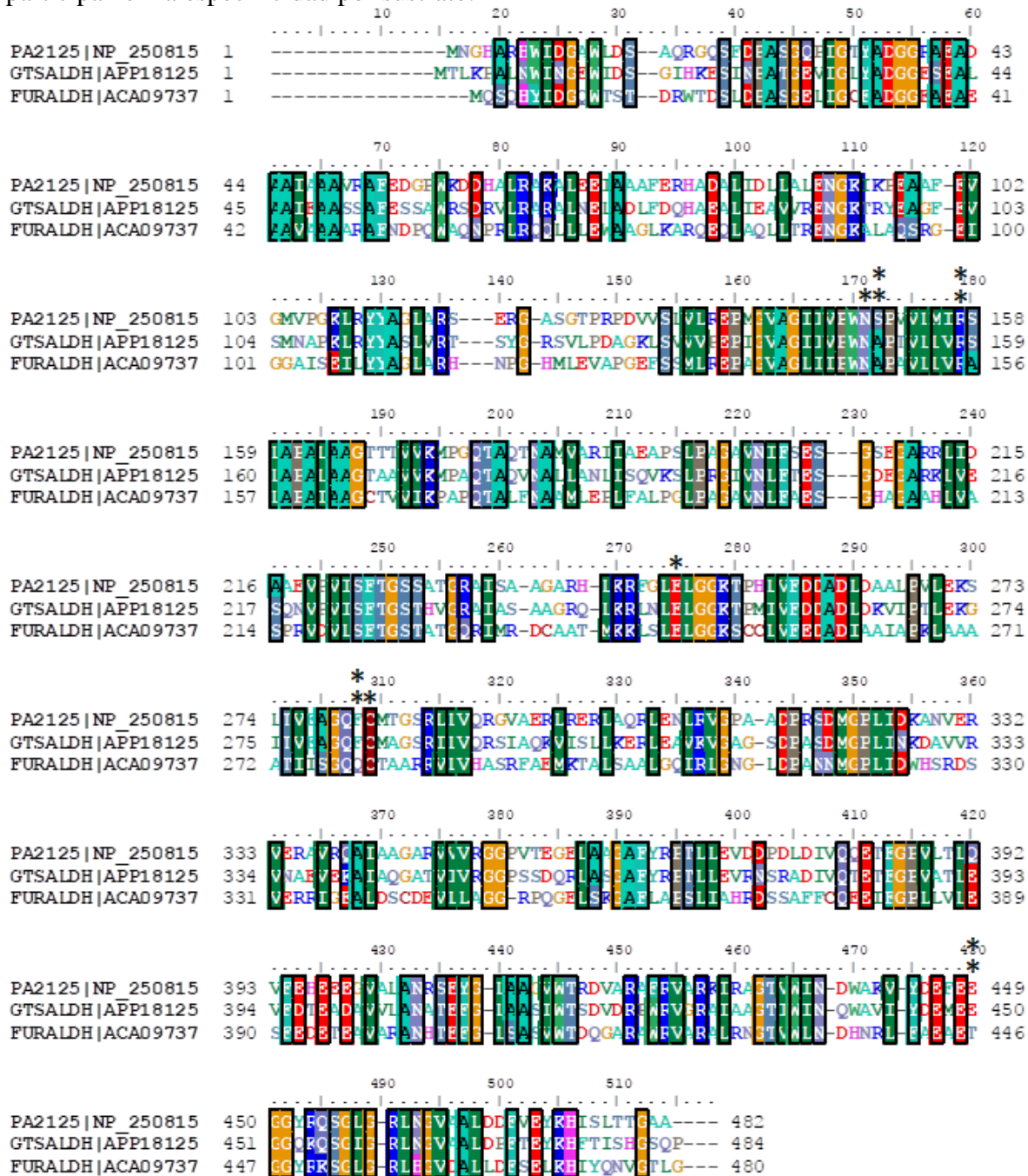
WP_049453296.1	Aldehído deshidrogenasa [ <i>Stenotrophomonas maltophilia</i> ]	47	67	469
WP_068257157.1	Aldehído deshidrogenasa [ <i>Janibacter corallicola</i> ]	47	66	469
WP_035713236.1	Aldehído deshidrogenasa [ <i>Azorhizobium doebereineriae</i> ]	47	60	478
WP_090592155.1	Aldehído deshidrogenasa [ <i>Pelagibacterium luteolum</i> ]	47	61	468
WP_078201748.1	Aldehído deshidrogenasa [ <i>Cupriavidus necator</i> ]	47	60	477
ACA09737.1	PsfA [ <i>Pseudomonas putida</i> ]	47	63	477
WP_084969975.1	Aldehído deshidrogenasa [ <i>Pantoea calida</i> ]	47	66	470
WP_090597310.1	Aldehído deshidrogenasa [ <i>Pelagibacterium luteolum</i> ]	47	62	473
WP_064175071.1	Aldehído deshidrogenasa [ <i>Klebsiella pneumoniae</i> ]	46	66	470
WP_051749670.1	Aldehído deshidrogenasa [ <i>Sphingobium chlorophenicum</i> ]	46	65	471
WP_062607759.1	Aldehído deshidrogenasa [ <i>Caballeronia calidae</i> ]	46	60	478
WP_005862055.1	Aldehído deshidrogenasa [ <i>Sagittula stellata</i> ]	46	61	476
WP_066659999.1	Aldehído deshidrogenasa [ <i>Bordetella flabilis</i> ]	46	61	474
WP_086966373.1	Aldehído deshidrogenasa [ <i>Caballeronia glebae</i> ]	46	61	479
WP_067756450.1	Aldehído deshidrogenasa [ <i>Orrella dioscoreae</i> ]	46	61	482
WP_061147754.1	Aldehído deshidrogenasa [ <i>Caballeronia arvi</i> ]	46	61	478
WP_090441648.1	Aldehído deshidrogenasa [ <i>Pseudomonas benzenivorans</i> ]	46	63	477

WP_090045577.1	Aldehído deshidrogenasa [ <i>Lechevalieria fradiae</i> ]	46	64	449
WP_042444757.1	Aldehído deshidrogenasa [ <i>Azospirillum lipoferum</i> ]	46	59	476
WP_008328507.1	Aldehído deshidrogenasa [ <i>Maritimibacter alkaliphilus</i> ]	45	60	476
WP_088708905.1	Aldehído deshidrogenasa [ <i>Noviherbaspirillum denitrificans</i> ]	45	59	478
AIJ45734.1	succinate-semialdehyde dehydrogenase [ <i>Comamonas testosteroni TK102</i> ]	45	59	474
WP_066703930.1	Aldehído deshidrogenasa [ <i>Curvibacter delicatus</i> ]	45	59	482
SMF89477.1	Acyl-CoA reductase [ <i>Azospirillum oryzae</i> ]	45	59	479
WP_088605029.1	Aldehído deshidrogenasa [ <i>Candidimonas nitroreducens</i> ]	45	60	475
WP_028165735.1	Aldehído deshidrogenasa [ <i>Bradyrhizobium elkanii</i> ]	45	60	482
WP_089883907.1	Aldehído deshidrogenasa [ <i>Citriimonas salinaria</i> ]	44	59	480
WP_076694628.1	Aldehído deshidrogenasa [ <i>Pelagibaca abyssii</i> ]	44	59	478
WP_066344197.1	Aldehído deshidrogenasa [ <i>Bordetella bronchialis</i> ]	44	59	481
AJE49484.1	Aldehído deshidrogenasa (plasmid) [ <i>Celeribacter indicus</i> ]	44	61	406
WP_028604809.1	Aldehído deshidrogenasa [ <i>Ottowia thiooxydans</i> ]	44	58	473
WP_067287002.1	Aldehído deshidrogenasa [ <i>Marinobacterium profundum</i> ]	44	59	479
OHC09307.1	Aldehído deshidrogenasa [ <i>Polynucleobacter sp. GWA2_45_21</i> ]	43	58	479
WP_043872375.1	Aldehído deshidrogenasa [ <i>Celeribacter indicus</i> ]	43	59	478

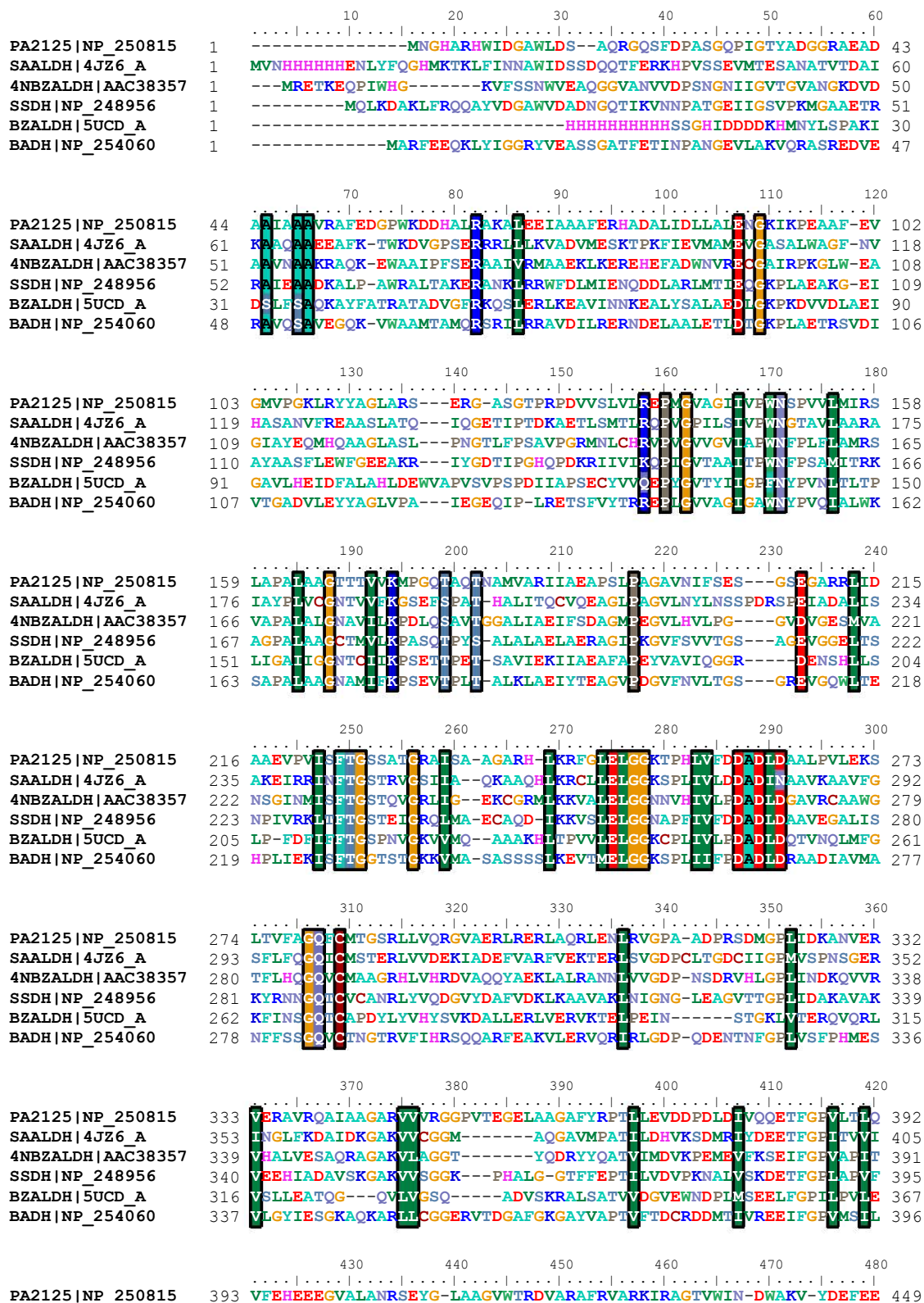
WP_089727394.1	Aldehído deshidrogenasa [ <i>Halomonas muralis</i> ]	43	59	478
WP_062308886.1	Aldehído deshidrogenasa [ <i>Polynucleobacter sinensis</i> ]	43	58	481
ODS72760.1	Aldehído deshidrogenasa [ <i>Bordetella sp.</i> SCN 67-23]	43	60	454
WP_028672207.1	Aldehído deshidrogenasa [ <i>Saccharospirillum impatiens</i> ]	43	58	473
WP_090270724.1	Aldehído deshidrogenasa [ <i>Litorimicrobium taeanense</i> ]	43	61	480
WP_043375884.1	Aldehído deshidrogenasa [ <i>Comamonas testosteroni</i> ]	43	60	475
WP_017523433.1	Aldehído deshidrogenasa [ <i>Pusillimonas noertemanni</i> ]	42	61	485
ALP00403.1	PutA [ <i>Pseudomonas putida</i> ]	42	58	475
WP_012187090.1	Aldehído deshidrogenasa [ <i>Dinoroseobacter shibae</i> ]	42	57	476
WP_089400633.1	Aldehído deshidrogenasa [ <i>Noviherbaspirillum humi</i> ]	42	60	483
WP_091317136.1	Betaína aldehído deshidrogenasa [ <i>Amycolatopsis tolypomycina</i> ]	42	58	476

---

**Anexo 2. Alineamientos de secuencias de aminoácidos entre PA2125, gentaldehído deshidrogenasa (GTSALDH) y furfural deshidrogenasa (FURALDH).** Los colores indican la naturaleza química de las cadenas laterales de los residuos. Residuos con un asterisco corresponden a los residuos catalíticos y residuos con dos asteriscos residuos que participan en la especificidad por sustrato.



**Anexo 3. Alineamientos de secuencias de aminoácidos entre PA2125 y enzimas que presentan actividad con aldehídos aromáticos derivados del BZAL, con succínico semialdehído, y BAL. Salicilaldehído deshidrogenasa (SAALDH; PDB 4JZ6); 4-NITROBZAL deshidrogenasa (4NBZALDH; AAC38357), succínico semialdehído deshidrogenasa (SSDH; NP\_248956), BZAL deshidrogenasa (BZALDH; PDB: 5UCD). Betaína aldehído deshidrogenasa de *P. aeruginosa* (BADH; NP\_254060).**



		430	440	450	460	470	480	
		..... ..... ..... ..... ..... ..... ..... .....						
PA2125 NP_250815	393	VFEHEEEGV	ALANRSEYG-	LAAGVWTRD	VARAFRVARKI	RAGTVWIN-	DWAKV-YDEFEE	449
SAALDH 4JZ6_A	406	RCKGEAEAIR	LANDSVYG-	LSSGVFGRD	INRALRVGMSI	EYGCVIN-	GSTVQNEAQAPY	463
4NBZALDH AAC38357	392	VFDSIEEAIE	LANSSEYG-	LAASIH	TTALATGLD	IAKRLNTGMV	HIN-DQPINCEPHVPF	449
SSDH NP_248956	396	RFKDEAEVI	AMSNDTEFG-	LASYFYARD	LARVFRVAEQ	LEYGMVGIN-	TGLIS-NEVAPF	452
BZALDH 5UCD_A	368	FDSVRTAIDQ	VNKHHPKPL	AVYVFGKDM	VAKGIINQIQ	SGDAQVNGV	MHLHAFSPYLPFG	427
BADH NP_254060	397	VYDDEDEAIR	RANDTEYG-	LAAGVVTQD	LARAHRAIHR	LEAGICWIN-	TWGES-PAEMPV	453
		490	500	510				
		..... ..... ..... ..... ..... .....						
PA2125 NP_250815	450	GGYRQSGLG-	RLNGVAALDD	FVEYKHISL	TTGAA----			482
SAALDH 4JZ6_A	464	GSTKNTGYG-	RFDGRAVIDE	FTELKWLTE	EPFEQQYPF			500
4NBZALDH AAC38357	450	GGMGSSGSGG	RFGGLASIDE	F	TQSQWISMVE	KSASYPF		487
SSDH NP_248956	453	GSIKASGLG-	REGSKYCI	EDYLEIK--	YLCLGGI----			483
BZALDH 5UCD_A	428	GIGASGMG-	EYHGHFSYL	TFTHKKS	VRIVP-----			456
BADH NP_254060	454	GSYKQSGVG-	RENGLTTLA	HYTRIKSVQ	VELGDYASVF			490

**Anexo 4. Alineamiento de la secuencia de nucleótidos obtenida del gen clonado a partir del DNA genómico de *P. aeruginosa* con la secuencia del gen *pa2125* depositada en la base datos del NCBI.** Las bases que codifican la etiqueta de histidinas en el gen clonado están enmarcadas en rojo, las de la secuencia de reconocimiento de la endonucleasa NdeI enmarcadas en verde y las de secuencia de reconocimiento de HindIII enmarcadas en azul).

Score	Expect	Identities	Gaps	Strand
2776 bits(1503)	0.0	1503/1503(100%)	0/1503(0%)	Plus/Plus
Query 1	catcatcatcatcatca	CAGCAGCGGCCTGGTGCCGCGCGGCAGCCATATGAACGGACAC		60
Sbjct 1	CATCATCATCATCATCA	CAGCAGCGGCCTGGTGCCGCGCGGCAGCCATATGAACGGACAC		60
Query 61	GCCAGACACTGGATCGACGGCGCATGGCTCGATTCCGCGCAGCGGGGACAATCCTTCGAC			120
Sbjct 61	GCCAGACACTGGATCGACGGCGCATGGCTCGATTCCGCGCAGCGGGGACAATCCTTCGAC			120
Query 121	CCCGCCAGCGGCCAGCCGATCGGCACCTACGCCGACGGCGGGCGGGCCGAGGCCGACGCG			180
Sbjct 121	CCCGCCAGCGGCCAGCCGATCGGCACCTACGCCGACGGCGGGCGGGCCGAGGCCGACGCG			180
Query 181	GCGATCGCCGCGCTGTGCGCGCCTTCGAAGATGGTCCCTGGAAGGACGACCACGCGCTG			240
Sbjct 181	GCGATCGCCGCGCTGTGCGCGCCTTCGAAGATGGTCCCTGGAAGGACGACCACGCGCTG			240
Query 241	CGCGCGAAAGCCCTGGAGGAAATCGCCGCGGCCTTCGAGCGGCATGCCGATGCGCTCATC			300
Sbjct 241	CGCGCGAAAGCCCTGGAGGAAATCGCCGCGGCCTTCGAGCGGCATGCCGATGCGCTCATC			300
Query 301	GACCTGCTTGCCCTGGAGAACGGCAAGATCAAGCCCGAGGCTGCCTTCGAGGTGGGCATG			360
Sbjct 301	GACCTGCTTGCCCTGGAGAACGGCAAGATCAAGCCCGAGGCTGCCTTCGAGGTGGGCATG			360
Query 361	GTGCCGGGCAAGCTGCGCTACTACGCCGGCCTGGCGCGCAGCGAACGGGGCGCCAGCGGC			420
Sbjct 361	GTGCCGGGCAAGCTGCGCTACTACGCCGGCCTGGCGCGCAGCGAACGGGGCGCCAGCGGC			420
Query 421	ACGCCGCGCCCCGACGTGGTGTGCTGGTGTGCTGCGCGAGCCGATGGGCGTCGCCGGCATC			480
Sbjct 421	ACGCCGCGCCCCGACGTGGTGTGCTGGTGTGCTGCGCGAGCCGATGGGCGTCGCCGGCATC			480
Query 481	ATCGTGCCGTGGAATTCACCGGTGGTGTGATGATCCGCTCCCTGGCGCCGGCGCTGGCC			540
Sbjct 481	ATCGTGCCGTGGAATTCACCGGTGGTGTGATGATCCGCTCCCTGGCGCCGGCGCTGGCC			540
Query 541	GCCGGTACCACCACGGTGGTGAAGATGCCGGGGCAGACCGCGCAGACCAACGCCATGGTC			600
Sbjct 541	GCCGGTACCACCACGGTGGTGAAGATGCCGGGGCAGACCGCGCAGACCAACGCCATGGTC			600
Query 601	GCGCGGATCATCGCCGAAGCGCCGTGCTGCCTGCCGGGGCGGTGAACATCTTCAGCGAA			660
Sbjct 601	GCGCGGATCATCGCCGAAGCGCCGTGCTGCCTGCCGGGGCGGTGAACATCTTCAGCGAA			660
Query 661	TCGGGTTCCGAAGGCGCCCGGCGATTGATCGACGCGGCCGAGGTGCCGGTGTATCAGCTTC			720
Sbjct 661	TCGGGTTCCGAAGGCGCCCGGCGATTGATCGACGCGGCCGAGGTGCCGGTGTATCAGCTTC			720
Query 721	ACCGGCTCGAGCGCCACCGGCCGAGCGATTTCCGCCGCGGGCGCGGCACCTGAAGCGC			780
Sbjct 721	ACCGGCTCGAGCGCCACCGGCCGAGCGATTTCCGCCGCGGGCGCGGCACCTGAAGCGC			780
Query 781	TTCGGCCTGGAAGTGGGCGGCAAGACACCGCACCTGGTGTTCGACGACCGACCTGGAC			840
Sbjct 781	TTCGGCCTGGAAGTGGGCGGCAAGACACCGCACCTGGTGTTCGACGACCGACCTGGAC			840



```

Query 841  GCCGCCTTGCCCGTGTGGAGAAGTCGCTGACGGTGTTCGCCGGGCAGTTCTGCATGACC 900
          |||
Sbjct 841  GCCGCCTTGCCCGTGTGGAGAAGTCGCTGACGGTGTTCGCCGGGCAGTTCTGCATGACC 900

Query 901  GGTTCGCGCCTGCTGGTGCAGCGCGCGTTCGCCGAACGGCTACGCGAACGCCTGGCGCAG 960
          |||
Sbjct 901  GGTTCGCGCCTGCTGGTGCAGCGCGCGTTCGCCGAACGGCTACGCGAACGCCTGGCGCAG 960

Query 961  CGCCTGGAAAACCTGCGCGTGGGGCCGGCGGGACCCGCGCAGCGACATGGGGCCGCTG 1020
          |||
Sbjct 961  CGCCTGGAAAACCTGCGCGTGGGGCCGGCGGGACCCGCGCAGCGACATGGGGCCGCTG 1020

Query 1021 ATCGACAAGGCCAACGTCGAGCGCGTCGAGCGAGCCGTACGCCAGGCCATCGCCGCCGGC 1080
          |||
Sbjct 1021 ATCGACAAGGCCAACGTCGAGCGCGTCGAGCGAGCCGTACGCCAGGCCATCGCCGCCGGC 1080

Query 1081 GCGCGGGTGGTGGTCCGCGCGCGCCCGTAACCGAGGGCGAGCTGGCGCCGGCGCCTTC 1140
          |||
Sbjct 1081 GCGCGGGTGGTGGTCCGCGCGCGCCCGTAACCGAGGGCGAGCTGGCGCCGGCGCCTTC 1140

Query 1141 TACCGTCCGACCCTGCTGGAGGTCGATGATCCGGACCTGGACATCGTGAACAGGAAAAC 1200
          |||
Sbjct 1141 TACCGTCCGACCCTGCTGGAGGTCGATGATCCGGACCTGGACATCGTGAACAGGAAAAC 1200

Query 1201 TTCGGACCGGTGCTGACCCTGCAGGTATTTCGAGCACGAGGAGGAGGGCGTCGCCCTGGCC 1260
          |||
Sbjct 1201 TTCGGACCGGTGCTGACCCTGCAGGTATTTCGAGCACGAGGAGGAGGGCGTCGCCCTGGCC 1260

Query 1261 AACCGCAGCGAGTACGGGCTGGCGGCCGGGGTCTGGACCCGCGATGTCGCGCGTGCCTTC 1320
          |||
Sbjct 1261 AACCGCAGCGAGTACGGGCTGGCGGCCGGGGTCTGGACCCGCGATGTCGCGCGTGCCTTC 1320

Query 1321 CGGGTGGCGCGGAAGATTTCGCGCCGGCACCGTGTGGATCAACGACTGGGCCAAGGTCTAC 1380
          |||
Sbjct 1321 CGGGTGGCGCGGAAGATTTCGCGCCGGCACCGTGTGGATCAACGACTGGGCCAAGGTCTAC 1380

Query 1381 GACGAATTCGAGGAGGGCGGCTACCGCCAGTCCGGCCTCGGCCGGCTGAACGGGGTTCGCC 1440
          |||
Sbjct 1381 GACGAATTCGAGGAGGGCGGCTACCGCCAGTCCGGCCTCGGCCGGCTGAACGGGGTTCGCC 1440

Query 1441 GCCCTGGACGACTTCGTCGAATACAAGCACATCAGCCTGACCACCGGCGCGGCCTGAAG 1500
          |||
Sbjct 1441 GCCCTGGACGACTTCGTCGAATACAAGCACATCAGCCTGACCACCGGCGCGGCCTGAAG 1500

Query 1501 CTT 1503
          |||
Sbjct 1501 CTT 1503

```

**Anexo 5. Resultados de la evaluación de estructuras cristalográficas como posibles moldes para la obtención del modelo por homología de PA2125. Se usó el método HHSENSOR, (Söding *et al.*, 2006) \* INDEX, puntuaciones HMM como un porcentaje de la longitud de la secuencia alineada**

No	PDB hit			Fitness	E-value (Probability)	Score(±SS):::columns	Coverage		Quality	
	code	chain	Description (short)				Query::Template	HMM score	INDEX*	Completeness
1	4QF6	A	Aldehyde_dehydrogenase;	100	1x10 <sup>-101</sup> (P=4x10 <sup>-106</sup> )	803(±58.7):::475	1→480::16→494	494	3%	98.56%
2	4PXL	A	Cytosolic_aldehyde_dehy	100	5x10 <sup>-101</sup> (P=1x10 <sup>-105</sup> )	801.4(±62):::474	4→482::36→514	517	8%	98.56%
3	4YWE	A	Putative_aldehyde_dehyd	100	1x10 <sup>-101</sup> (P=4x10 <sup>-106</sup> )	801.3(±54.1):::466	6→480::13→487	487	3%	97.73%
4	3EK1	A	Aldehyde_dehydrogenase;_	100	9x10 <sup>-101</sup> (P=2x10 <sup>-105</sup> )	797.8(±57.1):::468	4→480::32→503	504	6%	97.11%
5	1A4S	A	ALDH;_betaine_aldehyde_	100	8x10 <sup>-101</sup> (P=2x10 <sup>-105</sup> )	798.7(±56):::471	4→482::19→499	503	5%	98.97%
6	4QTO	A	Betaine_aldehyde_dehydr...	100	5x10 <sup>-100</sup> (P=1x10 <sup>-104</sup> )	795(±61.2):::469	4→480::32→511	520	9%	98.76%
7	4O5H	A	Phenylacetaldehyde_dehy	100	2x10 <sup>-101</sup> (P=7x10 <sup>-106</sup> )	804.1(±50.3):::473	1→480::27→511	511	7%	99.79%
8	4PZ2	A	Zmaldh;_oxidoreductase;_	100	4x10 <sup>-100</sup> (P=1x10 <sup>-104</sup> )	797.3(±57.9):::471	4→481::50→530	534	12%	98.97%
9	2D4E	A	5-carboxymethyl-2-hydro...	100	8x10 <sup>-100</sup> (P=2x10 <sup>-104</sup> )	793.5(±57.7):::474	1→482::23→508	515	7%	100.00%
10	1O04	A	Aldehyde_dehydrogenase...	100	4x10 <sup>-100</sup> (P=1x10 <sup>-104</sup> )	794(±53):::471	4→481::19→497	500	5%	98.56%
11	3IFG	A	Succinate-semialdehyde_	100	8x10 <sup>-100</sup> (P=2x10 <sup>-104</sup> )	790(±55.3):::469	4→479::12→484	484	2%	97.32%
12	1WND	A	Putative_betaine_aldehy	100	3x10 <sup>-100</sup> (P=9x10 <sup>-105</sup> )	792.7(±51.5):::468	4→479::23→495	495	4%	97.32%

**Anexo 6. Puntuaciones RD.HMM de secuencias completas recuperadas por el modelo de Markov oculto de la base de datos de secuencias nr NCBI.**

PDB id	HitNº	E-value	score	bias	Sequence found	Description
4EK1	1	6.80E-109	375.1	0	gi 513137405 pdb 4JWS A	Chain A, Crystal Structure Of Cy
4MPB	1	4.80E-149	508	1.7	gi 508773329 gb EOY20585.1	Betaine aldehyde dehydrogenase 1
4MPB	25302	3.80E-048	174.6	1.9	gi 15597321 ref NP_250815.1	aldehyde dehydrogenase [Pseudomo
4PXL	1	3.20E-142	485.6	5	gi 342883511 gb EGU83982.1	hypothetical protein FOXB_05503
4PXL	33924	3.30E-032	122.3	3	gi 15597321 ref NP_250815.1	aldehyde dehydrogenase [Pseudomo
4PZ2	1	4.80E-136	465.1	0.9	gi 413943806 gb AFW76455.1	hypothetical protein ZEAMMB73_71
4PZ2	28598	8.50E-039	143.8	7.3	gi 15597321 ref NP_250815.1	aldehyde dehydrogenase [Pseudomo
4Q92	1	4.70E-122	418.8	0.9	gi 4574120 gb AAD23900.1 AF00	glycine betaine aldehyde dehydro
4Q92	23223	2.90E-042	155.3	2.4	gi 15597321 ref NP_250815.1	aldehyde dehydrogenase [Pseudomo
4QF6	1	2.90E-132	452.5	5.9	gi 342883511 gb EGU83982.1	hypothetical protein FOXB_05503
4QF6	26900	9.20E-037	137.1	5.2	gi 15597321 ref NP_250815.1	aldehyde dehydrogenase [Pseudomo
4QTO	1	5.90E-135	461.6	0.6	gi 4574120 gb AAD23900.1 AF00	glycine betaine aldehyde dehydro
4QTO	24653	4.10E-045	164.8	2.1	gi 15597321 ref NP_250815.1	aldehyde dehydrogenase [Pseudomo
4YWE	1	1.30E-130	447.1	14.7	gi 206560962 ref YP_002231727.	putative aldehyde dehydrogenase
4YWE	22979	8.70E-033	124.1	7.3	gi 15597321 ref NP_250815.1	aldehyde dehydrogenase [Pseudomo
4ZWL	1	2.60E-123	422.8	0.2	gi 379796923 ref YP_005326924.	putative betaine aldehyde dehydr
4ZWL	24284	7.50E-043	157	0.2	gi 15597321 ref NP_250815.1	aldehyde dehydrogenase [Pseudomo

**Anexo 7. Resultados de los experimentos de docking con aldehídos aromáticos derivados del BZAL.** La columna "puntuación" muestra la afinidad de unión predicha (medida como energía de unión en kcal/mol). Las columnas "RMSD u.b." (upper bound: alta unión por sus siglas en inglés) y "rmsd l.b." (lower bound: baja unión por sus siglas en inglés) son dos variantes de los valores de desviación de la raíz cuadrada media (rmsd) que miden la coincidencia relativa al mejor modo de enlace utilizando solo átomos pesados móviles.

#### BZAL

Modo de unión	Puntuación (kcal/mol)	RMSD l.b.	RMSD u.p.
1	-4.1	0.0	3.1
2	-4.1	2.153	3.2
3	-3.9	0.556	3.3
4	-3.6	2.02	3.4
5	-3.6	1.466	3.5
6	-3.4	1.992	3.6
7	-3.4	3.746	3.7
8	-3.4	3.74	3.8
9	-2.9	3.148	3.9
10	-2.8	2.446	3.10

#### 4-METOXIBZAL

Modo de unión	Score (kcal/mol)	RMSD l.b.	RMSD u.p.
1	-3.8	0.0	0.0
2	-3.7	2.349	5.08
3	-3.6	2.248	4.929
4	-3.6	0.356	1.53
5	-3.6	0.853	4.28
6	-3.2	2.24	3.461
7	-3.1	2.681	3.049
8	-1.3	1.87	2.099
9	-1.2	5.603	5.943

#### SAAL

Modo de unión	Score (kcal/mol)	RMSD l.b.	RMSD u.p.
1	-4.3	0.0	0.0
2	-4.2	1.173	1.409
3	-4.1	0.889	2.401

4	-4.1	1.957	2.29
5	-3.8	1.992	2.56
6	-3.8	2.311	3.831
7	-3.2	2.094	3.795

### 3-NITROBZAL

Modo de unión	Score (kcal/mol)	RMSD l.b.	RMSD u.p.
1	-3.6	0.0	0.0
2	-3.4	1.760	4.104
3	-2.6	3.235	4.685
4	-2.4	2.706	3.315
5	-1.7	2.948	4.325
6	-0.8	3.534	4.550
7	-0.8	3.215	3.745
8	-0.7	2.884	3.676

### 4-NITROBZAL

Modo de unión	Score (kcal/mol)	RMSD l.b.	RMSD u.p.
1	-3.8	0.0	0.0
2	-3.8	0.048	1.715
3	-3.4	4.193	5.608
4	-3.3	1.886	4.908
5	-2.7	2.979	4.795
6	-2.4	4.020	4.898
7	-2.4	1.412	1.75
8	-2.0	3.857	4.815
9	-1.8	3.662	4.626
10	-1.6	5.359	6.073

### Furfural

Modo de unión	Score (kcal/mol)	RMSD l.b.	RMSD u.p.
1	-3.5	0.0	0.0
2	-3.4	1.493	3.362
3	-3.3	1.79	3.418
4	-3.2	1.395	3.176
5	-3.2	3.806	5.192
6	-3.1	1.282	3.270
7	-3.1	1.455	1.679
8	-3.1	2.300	3.363
9	-3.0	1.000	1.428

---

10	-2.7	3.265	4.005
----	------	-------	-------

---

Cinamaldehído

---

Modo de unión	Score (kcal/mol)	RMSD l.b.	RMSD u.p.
1	-4.5	0.0	0.0
2	-2.0	5.906	6.898
3	-1.9	3.938	4.627
4	-1.9	6.354	6.858

---

TMABAL

---

Modo de unión	Score (kcal/mol)	RMSD l.b.	RMSD u.p.
1	-3.2	0.0	0.0
2	-3.2	0.157	1.439
3	-2.7	1.408	2.086
4	-2.7	2.996	4.117
5	-2.6	3.3	4.419
6	-2.5	2.953	4.499
7	-2.1	2.608	4.281
8	-1.9	1.951	2.637
9	-1.0	2.984	4.569
10	-0.5	2.832	4.889

---

2-METOXIBZAL

---

Modo de unión	Score (kcal/mol)	RMSD l.b.	RMSD u.p.
1	-3.8	0.0	0.0
2	-3.7	0.812	2.909
3	-3.6	2.154	3.512
4	-3.5	1.939	3.264
5	-3.2	2.051	4.135
6	-3.0	1.286	2.952
7	-2.9	1.635	3.103
8	-2.5	2.863	3.802
9	-2.5	1.431	3.253
10	-2.1	3.923	4.941

---

---

BAL

Modo de unión	Score (kcal/mol)	RMSD l.b.	RMSD u.p.
1	-2.7	0.0	0.0
2	-2.7	2.269	3.083
3	-2.7	0.156	1.651
4	-2.7	0.115	1.608
5	-2.6	2.166	3.208
6	-2.2	2.414	3.388
7	-1.7	1.455	1.606
8	-1.0	2.367	3.412

LACTALDEHÍDO

Modo de unión	Score (kcal/mol)	RMSD l.b.	RMSD u.p.
1	-2.7	0.0	0.0
2	-2.6	1.429	1.125
3	-2.5	2.564	3.556
4	-2.5	2.358	3.367
5	-2.5	2.42	2.859
6	-2.4	1.731	1.892
7	-2.4	0.785	2.808
8	-2.4	2.243	2.985
9	-2.2	2.536	3.151
10	-1.9	5.062	5.71

## Anexo 8. Comparación contextos genómicos de enzimas de la familia cd07120

