



**UNIVERSIDAD NACIONAL AUTÓNOMA DE
MÉXICO**

**FACULTAD DE ESTUDIOS SUPERIORES
IZTACALA**

Análisis del gen *ectC* de *Halomonas spp.* del Valle de
Zapotitlán Salinas, Puebla.

TESIS

QUE PARA OBTENER EL TÍTULO DE

BIÓLOGO

PRESENTA

EDGAR GEOVANNI HURTADO MONZÓN

DIRECTOR DE TESIS

DRA. MARTHA MARTÍNEZ GARCÍA



Los Reyes Iztacala, Edo. de México, 2015



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



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.

DEDICATORIAS

A mis padres, Josefina y Ramón: por su amor, trabajo y sacrificios en todos estos años, gracias a ustedes he logrado llegar hasta aquí y convertirme en lo que soy. Por apoyarme en cada decisión que he tomado y siempre estar a mi lado para darme su consejo. Es todo un privilegio ser su hijo. Son los mejores padres que alguien podría tener.

A mi hermana, Arianna: por estar siempre a mi lado apoyándome y dándome consejos, por ser mi cómplice y por todos esos momentos buenos y malos, porque siempre serás un modelo a seguir para mí y algún día lograré alcanzarte, sabes que te amo.

A mi familia: por cada momento compartido, porque siempre estarán para apoyarme y porque simplemente son la mejor familia, los amo a todos.

A Marco e Ilse: por las pláticas y los momentos que vivimos en la carrera, por aguantar mis malos momentos y porque siempre formaran parte de mí.

A mis amigos: que siguen estando para mí a pesar de los años y la distancia, por escucharme y porque simplemente nos apoyamos sin importar nada, gracias.

A mi asesora Dra. Martha: por creer en mí, por sus consejos y enseñanzas, porque además de ser una gran asesora es una gran amiga, que me ha enseñado a ser perseverante y a confiar en lo mucho que puedo lograr, y que la diversión puede ir de la mano con el trabajo, me llevo mucho de usted, por todo gracias.

AGRADECIMIENTOS

Dra. Martha Martínez García, gracias por aceptarme en el laboratorio, por orientarme en la realización de este trabajo y por todo el apoyo.

M. en C. Alejandro Cruz Monsalvo Reyes, gracias por el apoyo y la ayuda en los experimentos de este trabajo, por compartirme sus conocimientos y los consejos brindados.

A mis sinodales Dr. Jorge Eduardo Campos Contreras, Dr. Rodolfo de la Torre Almaraz, Dra. María Graciela Molina González, por brindarme su tiempo, conocimientos y orientación que permitieron hacer de esta tesis un mejor trabajo.

A mis profesores, por compartirme sus conocimientos, por aquellos que se preocuparon por que aprendiera de verdad y porque sin ellos no podría haber concluido esta maravillosa carrera.

A esta institución, por brindarme el tiempo y los recursos para estudiar esta hermosa carrera.

ÍNDICE

RESUMEN.....	1
INTRODUCCIÓN.....	3
OBJETIVO GENERAL.....	11
Objetivos particulares.....	11
MATERIAL Y MÉTODOS.....	12
Material biológico.....	12
Extracción de DNA.....	12
Obtención del amplicón de la región del <i>ectC</i>	12
Secuenciación del gen.....	13
Análisis bioinformático.....	13
Dominios Conservados.....	14
Estructura en 3D.....	14
RESULTADOS.....	15
DISCUSIÓN.....	27
CONCLUSIONES.....	33
LITERATURA CITADA.....	35
ANEXO 1: Nomenclatura de cepas.....	43
ANEXO 2: Modelado en 3D.....	44
ANEXO 3: Símbolos de tres y una letra para cada aminoácidos.....	46

ÍNDICE DE FIGURAS

Figura 1	Síntesis de ectoína	7
Figura 2	Patrón electroforético con bandas inespecíficas.....	15
Figura 3	Amplicón de <i>ectC</i> ~300 pb	16
Figura 4	Alineamiento de secuencias	18
Figura 5	Dominios conservados de ectoína sintetasa	23
Figura 6	Azúcar fosfato isomerasa putativa	25
Figura 7	Quercitina 2,3 dioxigenasa	25
Figura 8	Cupina 2 proteína con un dominio conservado de barril.....	26

ÍNDICE DE CUADROS

Cuadro 1	BLAST con las secuencias nucleotídicas de <i>ectC</i> de Zapotitlán Salinas.....	17
Cuadro 2	Índice de polimorfismo entre las cepas de <i>Halomonas</i> en la región de 225pb de <i>ectC</i>	19
Cuadro 3a	Cambios de codón entre las secuencias nucleotídicas de las cepas de Zapotitlán Salinas	20
Cuadro 3b	Cambios de codón entre las secuencias nucleotídicas de las cepas de Zapotitlán Salinas y la secuencia con número de acceso AB119659.....	21
Cuadro 4	Modelos en 3D de la proteína teórica de <i>ectC</i>	24

RESUMEN

Las *Halomonas* son bacterias halófilas moderadas que presentan un requerimiento de salinidad para su óptimo crecimiento, la característica que les permite crecer en altas concentraciones de sal es la acumulación de ectoína, un soluto compatible que mantiene el equilibrio osmótico. La síntesis de esta molécula es resultado de la acción de las enzimas codificadas por los genes *ectABC*, siendo el responsable final de la biosíntesis el gen *ectC*, conocido también como ectoína sintetasa. En este trabajo, se reportan las primeras secuencias parciales del gen *ectC* de *Halomonas spp.* de México, para lo cual se secuenció una región del gen de seis cepas previamente aisladas del Valle de Zapotitlán Salinas, Puebla. Al analizar las secuencias nucleotídicas se encontró un índice de polimorfismo mayor a 16 % con respecto a la secuencia con mayor similitud del GenBank, mientras que entre ellas sólo se presentaba un máximo de 4.4 % en los 225 pb analizados del gen *ectC*; es importante subrayar que en todas las secuencias comparadas se mantiene sin cambios una región de 20 pb. Al traducir *in silico* las secuencias nucleotídicas en aminoácidos, se observó que la mayoría de los cambios mantienen teóricamente las características de la proteína. Asimismo, se realizó la búsqueda de dominios funcionales en este gen y con el uso de la bioinformática se generó un modelo teórico de la proteína ectoína sintetasa. Por otro lado, la herramienta de dominios conservados arrojó que la proteína ectoína sintetasa EctC tiene un dominio que la cataloga dentro de la superfamilia de las Cupinas, este mismo resultado se obtuvo al momento de la búsqueda de los modelos, que clasificaron la secuencia obtenida dentro de la superfamilia de las Cupinas. Se ha reportado que

basados en el agrupamiento de estructura y secuencia los miembros de dicha familia reflejan un agrupamiento funcional en respuesta a la falta de agua disponible.

INTRODUCCIÓN

Las bacterias halófilas son organismos que se encuentran en condiciones adversas para la mayoría de los seres vivos, por su contenido de salinidad, conocidos como ambientes hipersalinos, los cuales presentan una baja actividad de agua, que es una medida de la disponibilidad de agua que, debido a la elevada cantidad de sales está muy limitada. Lo anterior, implica una adaptación de los organismos, ya que esta menor disponibilidad de agua, impide realizar los procesos metabólicos generales, la hidratación de proteínas y de ácidos nucleicos, para ello poseen características bioquímicas y metabólicas que les permiten vivir en hábitats poco comunes (Litchfield, 1998). Dichos ambientes están extendidos por todo el mundo, sin embargo, los que son extremadamente salinos son raros. La mayor parte de estos hábitats se encuentran en zonas calientes y secas (Ramírez *et al.*, 2004). Además de la elevada cantidad de sales, las temperaturas elevadas y la luz intensa, los ambientes extremadamente salinos se caracterizan por registrar un pH alcalino y la presencia de otros aceptores de electrones además del oxígeno. La adaptación a las condiciones antes mencionadas ha permitido identificar algunas características con gran potencial biotecnológico (Margesin *et al.*, 2001; Van den Burg, 2003).

Una de las características que permiten el desarrollo de las bacterias halófilas en ambientes altamente salinos, es el mecanismo más versátil desde el punto de vista de las variaciones de sal llamado *salt out*, que en respuesta al estrés osmótico, acumulan moléculas polares sencillas, provenientes de azúcares, alcoholes, aminoácidos u otros derivados, altamente solubles y sin carga o con carga neutra, llamados solutos

compatibles que mantienen el equilibrio osmótico y son compatibles con el metabolismo celular (Brown, 1976). Además, en el caso contrario de un choque hiposmótico, las células pueden restaurar el equilibrio osmótico mediante la liberación de osmólitos a través de sistemas de eflujo específicos (Wood *et al.*, 2001; Morbach y Krämer, 2002). La mayoría de las bacterias usan una gama de solutos para el balance osmótico, dependiendo principalmente de la duración del estrés osmótico, nivel de salinidad, disponibilidad de sustratos y osmólitos en el medio o la fuente de carbón, siendo la ectoína (ácido 2-metil-1,4,5,6-tetrahidro-4-pirimidincarboxílico) ,uno de los solutos compatibles identificados y más abundante en bacterias (Margesin *et al.*, 2001; Ramírez *et al.*, 2004; Roberts, 2005).

El género *Halomonas* es uno de los taxones bacterianos más representativos dentro de las bacterias halófilas moderadas. Se encuentra incluido en el phylum Proteobacteria, clase *Gamma-Proteobacteria*, familia Halomonadaceae (Arahal y Ventosa, 2006). En la actualidad *Halomonas* contiene más de 70 especies, 12 de ellas han sido descritas desde los 80`s (Quesada *et al.*, 1984; Llamas *et al.*, 2011). Debido a este alto número de especies la heterogeneidad fenotípica entre ellas es elevada (Mata *et al.*, 2002; Garritand *et al.*, 2005; Arahal y Ventosa, 2006) en cuanto al contenido G+C este género tiene un rango amplio de 52-68 mol % (Arahal *et al.*, 2002), teniendo en cuenta que generalmente, en el contenido de G+C entre miembros de un mismo género, no debe diferir más de un 10% (Owen y Pitcher, 1985). De manera generalizada, se puede decir que se trata en su mayoría de bacterias halófilas moderadas que presentan un requerimiento de sal

para su óptimo crecimiento que oscila entre un 3-15% (p/v). Las *Halomonas* presentan una morfología de bacilos o bacilo-cocos, Gram negativos, quimioheterótrofos con un metabolismo versátil, con algunas especies flageladas. La mayoría de las especies realizan respiración aerobia y un gran número de ellas también utilizan otros compuestos como últimos aceptores de electrones, mientras que otro grupo minoritario es capaz de realizar fermentaciones. Otra característica a tener en cuenta es que estas bacterias pueden crecer en una gran variedad de compuestos como única fuente de carbono y energía (Arahal *et al.*, 2007).

En otros aspectos, el género *Halomonas* ha cobrado gran importancia industrial desde la década de los 90's debido a la producción de ectoína, por las diversas aplicaciones biotecnológicas que se han encontrado para esta molécula, tales como agente protector contra el calor, procesos de congelación y descongelación (Pastor *et al.*, 2010), protector solar en piel, tolerancia a la sal en plantas de tabaco transformadas, así como la protección de las proteínas y los efectos estabilizadores en el DNA (Ramírez *et al.*, 2004). Así, el interés de la industria por la obtención de estos metabolitos ha provocado que genes de algunas especies de *Halomonas*, responsables de la síntesis de ectoína se patenten por empresas internacionales como Merk (2000) y Sebapharma (2004, 2014).

La ectoína se describió por primera vez en la bacteria purpúrea fotosintética *Halorhodospira halochloris* (anteriormente *Ectothiorhodospira halochloris*), de donde proviene el nombre de ectoína (Galinski *et al.*, 1985), aunque se encuentra ampliamente distribuida en

bacterias. La ectoína se forma por deshidratación intramolecular del ácido N- γ -acetildiaminobutírico (NADA), que da lugar a la formación del derivado pirimidínico. La ruta biosintética de la ectoína fue elucidada en *H. halochloris* y *H. elongata* por Peters *et al.* en 1990. La síntesis de esta molécula es resultado de la acción de las enzimas codificadas por los genes *ectABC* (Lo *et al.*, 2009), el gen *ectC* es el responsable final, aunque para la síntesis de ectoína se tienen diversos precursores; se inicia principalmente en la vía de los aminoácidos, comienza con la fosforilación del L-aspartato mediante la enzima aspartato cinasa, se obtiene L-aspartato-fosfato, continua con la síntesis de L-aspartato-beta-semialdehído catalizado por la L-aspartato-beta-semialdehído deshidrogenasa. A partir de este intermediario general, la ruta específica de síntesis de la ectoína ocurre en tres pasos enzimáticos. Primero, el L-aspartato-beta-semialdehído es transformado en L-ácido diaminobutírico mediante la enzima L-ácido diaminobutírico transaminasa o EctB, subsecuentemente es acetilado convirtiéndose en N- γ -ácido diaminobutírico por la enzima L-ácido diaminobutírico acetil transferasa o EctA, finalmente hay una condensación cíclica que da paso a la formación de ectoína mediante la enzima ectoína sintetasa o EctC (Figura 1). Se han aislado y caracterizado parcialmente las enzimas implicadas en dicha ruta en *H. elongata*: el L-ácido diaminobutírico transaminasa (EctB) con un peso molecular de 42 kDa, la DA enzima L-ácido diaminobutírico acetil transferasa (EctA) con un peso molecular de 19 kDa (Ono *et al.*, 1999) y la NADA deshidratasa, más conocida como ectoína sintasa -EctC- con un peso molecular de 16 kDa (Schubert *et al.*, 2007).

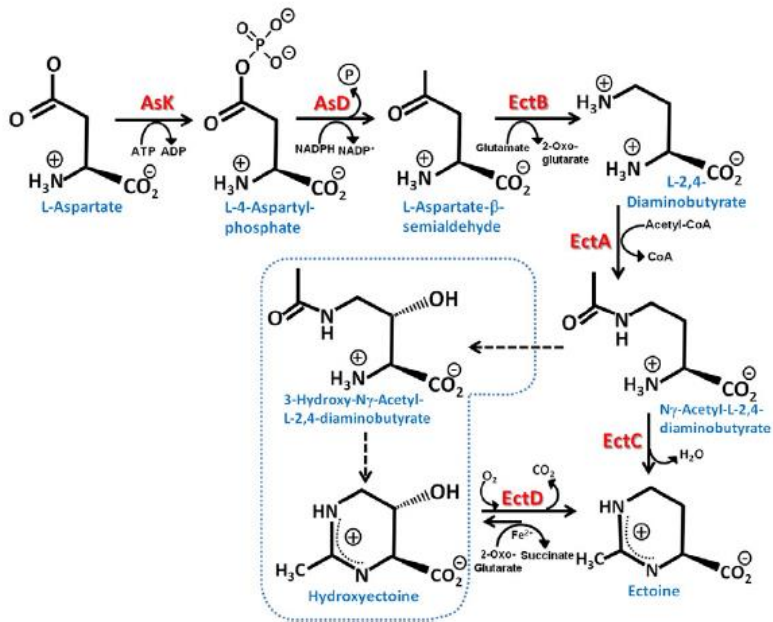


Figura 1. Síntesis de ectoína a partir del aminoácido Aspartato. Tomado de Pastor *et al.*, 2010.

La ectoína sintetasa al igual que otras enzimas responsables de procesos de regulación metabólica, presentan una estructura de dominios funcionales. Los dominios proteicos pueden ser considerados como diferentes unidades funcionales y/o estructurales. Estas dos clasificaciones coinciden a menudo como una cuestión de facto, ya que se ha encontrado que una cadena polipeptídica puede ser una unidad independiente de plegado y a su vez tener una función específica. Estos dominios a menudo se identifican como unidades recurrentes (secuencia o estructura), los cuales se pueden encontrar en diversos contextos (Marchler *et al.*, 2013).

En la evolución molecular, tales dominios pueden haber sido utilizados como bloques de construcción y haber sido recombinados en diferentes arreglos para modular la función de la proteína. Definimos dominios conservados como unidades recurrentes en la evolución molecular, las extensiones de los cuales se determinan por análisis de secuencia y estructura. Contienen patrones de secuencia o motivos, que permiten su detección en secuencias de polipéptidos. La distinción entre los dominios y motivos no es fuerte especialmente en el caso de unidades cortas repetitivas. Los motivos funcionales también están presentes fuera del alcance de los dominios estructuralmente conservados (Marchler *et al.*, 2013). Las diferencias en la secuencia de los monómeros de aminoácidos que integran macromoléculas homólogas, presentes en dos formas de vida, reflejan la distancia evolutiva existente entre ellas (Rodicio y Mendoza, 2004).

Para la caracterización estructural de las proteínas, se cuenta con técnicas experimentales, principalmente cristalografía de rayos X y resonancia magnética nuclear, mismas que proporcionan estructuras de alta resolución, aunque desafortunadamente sólo una pequeña parte de las proteínas se pueden caracterizar con estas técnicas, un ejemplo de ello es la estructura de la enzima EctD reportada en 2010 por Reuter *et al.*, mediante cristalografía de rayos X en *Virgibacillus salexigens*. El núcleo de la estructura de EctD consiste en una doble cadena de hélices β , formando una porción principal en el sitio activo de la enzima. La posición del ligando a hierro en el sitio activo es mediado por un sitio evolucionadamente conservado, motivo de unión de 2-His-1- hierro carboxilado. La estructura de ectD reportada puede servir como modelo

para otros estudios bioquímicos o estructurales de la familia de la enzima. Para una gran parte de la fracción de secuencias cuya estructura no se puede determinar experimentalmente, los métodos computacionales de predicción de estructura – bioinformática- ofrecen información valiosa, útil para explicar gran parte de los aspectos funcionales que se pueden derivar del conocimiento estructural (Moya, 2003).

En cuanto al estudio del gen *ectC* (ectoína sintetasa) en bacterias del género *Halomonas*, la primera secuencia fue en la *Halomonas sp.* KS-3, realizada por Min-Lu *et al.*, 1993, quienes basados en la estrategia del diseño de primers degenerados a partir de la secuencia proteica de la enzima, clonaron el producto de PCR para su posterior secuenciación. Además, se ha demostrado que la ectoína es el soluto compatible predominante en la familia Halomonadaceae cuando crece en altas concentraciones de sal (Zanjirband *et al.*, 2008). Sin embargo, en México no se han realizado trabajos acerca del gen *ectC* de *Halomonas*, ni tampoco relacionados a los dominios conservados a nivel de secuencia nucleotídica ni de la proteína que traduce este gen.

Es evidente que las bacterias del género *Halomonas* poseen una gran importancia en diversos aspectos a estudiar, desde su condición taxonómica, papel ecológico y potencial biotecnológico, entonces la búsqueda de estas especies bacterianas representa un objetivo primordial para el entendimiento de los ambientes semiáridos salinos como la Reserva de la Biósfera de Tehuacán-Cuicatlán, zona donde se reporta una gran biodiversidad y altos índices de endemismo de plantas

vasculares y animales, no obstante, lo referente a los trabajos con microorganismos son muy escasos, por lo que la riqueza de especies se describe de forma limitada (CONABIO, 2013). En el Laboratorio de Bioquímica Molecular de la UBIPRO, FES Iztacala se aislaron cepas de bacterias halófilas del Valle de Zapotitlán Salinas, Puebla, que incluyen bacterias del género *Halomonas*.

OBJETIVO GENERAL

- Analizar el gen *ectC* en cepas de *Halomonas* spp aisladas del Valle de Zapotitlán Salinas, Puebla.

OBJETIVOS PARTICULARES

- Determinar la secuencia nucleotídica parcial del gen *ectC* en cepas de *Halomonas* aisladas del Valle de Zapotitlán Salinas, Puebla.
- Identificar los dominios conservados en la secuencia del gen *ectC* de las cepas de *Halomonas* de Zapotitlán Salinas, Puebla y contrastarlos con especies que presenten una arquitectura similar en el dominio estudiado.
- Generar un modelo en tercera dimensión de la proteína teórica traducida de la secuencia del gen *ectC*.

MATERIALES Y MÉTODO

Material Biológico

Se utilizaron seis cepas de bacterias del Valle de Zapotitlán Salinas, Puebla (Anexo 1) previamente aisladas en el Laboratorio de Bioquímica Molecular de la UBIPRO, FES Iztacala. Estas cepas se crecieron sobre agar nutritivo al 10% de NaCl y se mantuvieron al 7.5% de NaCl. Las cepas fueron puestas en crioconsecuación.

Extracción de DNA genómico

Las seis cepas bacterianas se crecieron por 48 h en agar nutritivo al 7.5% de NaCl. La extracción de DNA se realizó siguiendo las instrucciones descritas en el manual de usuario del kit “Ultra Clean[®] Plant DNA Isolation” marca MoBio. Posteriormente se verificó la calidad de DNA genómico por electroforesis en un gel de agarosa al 0.8% en una solución buffer TBE (Tris-Borato-EDTA pH8.5) 0.5X teñido en una solución de bromuro de etidio (BrEt).

Obtención del Amplicón del gen *ectC*

Se realizó mediante la PCR punto final con la enzima MyTaq[®] en un termociclador marca BIO RAD mod. T100, se usaron los primers (Forward: 5'-GGTAAYTGGGAYAGYACRC-3' y Reverse: 5'-GBGGHGTRAAKACRCADCC-3' donde, Y=C o T, R=A o G, B=C, G o T H=A, C o T K=G o T y D=A, G o T) (Zanjirband *et al.*, 2008) y 80 ng de DNA genómico bacteriano. Para la elección de la temperatura apropiada para la reacción se utilizó un gradiente de temperaturas desde los 53 –

58°C. La amplificación se comprobó por electroforesis en un gel de agarosa al 1.5 %, teñido con BrEt. Una vez determinada la temperatura de alineamiento a 54°C, se efectuó una segunda amplificación utilizando como DNA molde una fracción correspondiente a 300 pb del producto de PCR. Ésta última reacción se llevó a cabo siguiendo las condiciones de la enzima MyTaq®. Al término de la PCR se visualizó el amplicón por electroforesis en un gel de agarosa al 1.5 % teñido en BrEt

Secuenciación del gen

La reacción de secuencia del amplicón del gen *ectC* se llevó a cabo usando el kit de secuencia Big Dye Terminator versión 3.1, y se corrió en un equipo ABI / Hitachi 3130xl Genetic Analyzer de la Unidad de Biotecnología y Prototipos de la FES Iztacala.

Análisis bioinformático

Con la secuencia nucleotídica obtenida se realizó una búsqueda en el banco de genes del National Center for Biotechnology Information (NCBI, por sus siglas en inglés) mediante el algoritmo Basic Local Alignment Search Tool (BLAST, *id.*), se compararon las secuencias entre las bacterias del Valle de Zapotitlán Salinas, Puebla contra las reportadas en la base de datos de nucleótidos del GenBank para encontrar diferencias mediante un alineamiento múltiple en el programa Geneious 7.1.3; asimismo, se construyó una matriz de polimorfismo a nivel nucleotídico para esta región del gen.

Dominios Conservados

Se realizó una traducción a aminoácidos a partir de las secuencias nucleotídicas de la región *ectC* pertenecientes a cada una de las cepas del Valle de Zapotitlán Salinas, Puebla, mediante el programa Geneious 7.1.3. La secuencia consenso resultante se sometió a un BLAST en la base de datos de proteínas. Además, la secuencia proteica consenso se analizó con la herramienta de dominios conservados (Conserved Domains Database, CDD) del NCBI -este análisis es una variante del BLAST-.

Estructura en 3D

Para obtener un modelo teórico de la proteína se siguió el método descrito por Moya (2003), donde utiliza la base de datos de SWISS-MODEL perteneciente a BIOZENTRUM de la Universidad de Basel - método recomendado para secuencias menores a 150 residuos de aminoácidos- (Anexo 2).

RESULTADOS

Como resultado de la primera reacción de amplificación utilizando los primers degenerados y un gradiente desde 53-58°C para la temperatura de alineación, en el carril 1 colocó el marcador de peso molecular y del carril 2-9 se encuentran hasta 7 bandas inespecíficas (Figura 2).

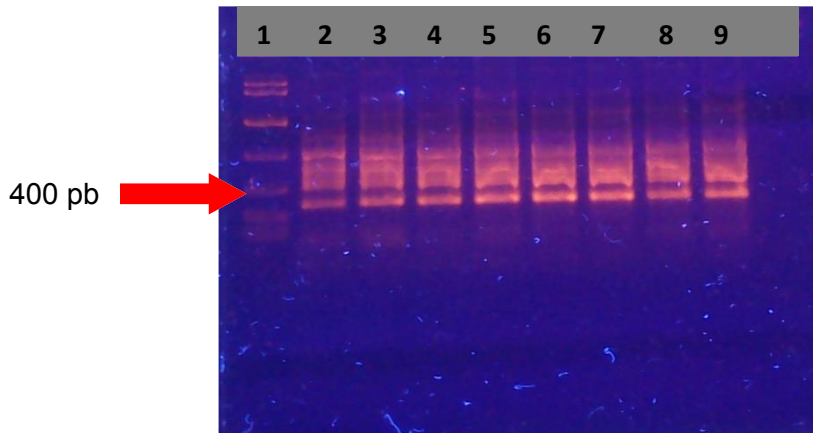


Figura 2. Patrón electroforético en gel de agarosa a 1.5%. Carril 1: escalera de peso molecular KAPA™ Express, la flecha señala 400 pb, carril 2-9 se observa un bandeo múltiple de hasta 7 bandas, cada carril muestra una temperatura diferente de alineamiento, de 58, 57.6, 57, 56.1, 54.9, 54, 53.3 y 53 °C respectivamente.

A las cepas que mostraron una banda de alrededor de 300pb, se purificaron y reamplificaron usando los mismos primers para obtener el amplicón, con la temperatura de alineamiento de 54°C.

Del total de los cultivos, sólo de 6 se logró la amplificación del fragmento de 300 pb después de la segunda amplificación como se muestra en la Figura 3.

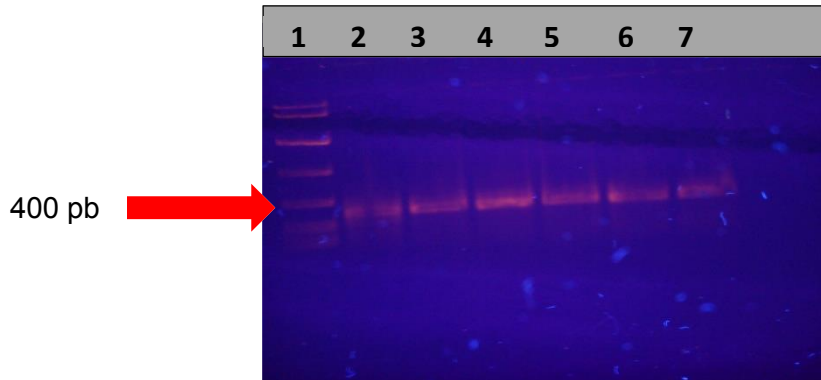


Figura 3. Amplicón obtenido con los primers para *ectC*. Las bandas alrededor de 300pb, carriles del 2-7 cepas Ia1, I- 2, Ia3, I- 6 , IVc9 y Vb4 respectivamente. Carril, escalera de peso molecular KAPA™ Express,

Los resultados del BLAST en la base de datos del GenBank usando las secuencias nucleotídicas de los amplicones de 300 pb se muestran en el Cuadro 1.

Cuadro1. Secuencias reportadas en el GenBank que presentan la mayor similitud con las secuencias de las cepas de *Halomonas* aisladas de Zapotitlán Salinas, Puebla.

Cepa	Descripción	Max Score	Total Score	Cobertura	E-value	Identidad	Núm de acceso
Ia1	Halomonas sp. QHL1 ectoine synthase (ectC) gene, complete cds	271	271	54%	5.00E-69	87%	JX312792.1
I-2	Halomonas sp. NT N60 ectC gene for ectoine synthasepartial cds, stain: NT N60	195	195	81%	4.00E-46	83%	AB196591.1
Ia3	Halomonas sp. QHL1 ectoine synthase (ectC) gene, complete cds	248	248	90%	3.00E-62	85%	JX312792.1
I-6	Halomonas sp. QHL1 ectoine synthase (ectC) gene, complete cds	163	163	87%	1.00E-36	79%	JX312792.1
IVc9	Halomonas sp. QHL1 ectoine synthase (ectC) gene, complete cds	250	250	80%	1.00E-62	85%	JX312792.1
Vb4	Halomonas sp. QHL1 ectoine synthase (ectC) gene, complete cds	239	239	85%	2.00E-59	84%	JX312792.1

Con la secuencia consenso resultado del alineamiento en el programa Geneious 7.1.3, se realizó un segundo BLAST, el cual mostró la mayor homología con la secuencia AB196591 perteneciente a *Halomonas sp. NT60 ectC* gene, misma que se utilizó como referencia para la edición del alineamiento. En total se alinearon 225pb, se observa una región de 20 pb (TTCGAAGGCACTGAGACACA), donde no presenta cambio alguno en los nucleótidos, esta condición se conserva en todas las secuencias analizadas (Figura 4).

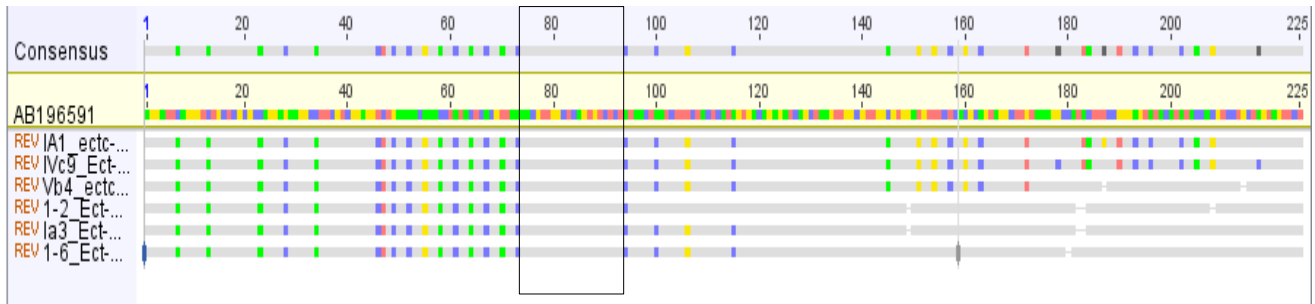


Figura 4. Alineamiento de las secuencias en el programa Geneious 7.1.3, los cuadros de color muestran los cambios. Se observa un sitio altamente conservado en la posición 74-94.

Por otro lado, en la secuencia de 225 pb, se encontraron hasta 10 cambios entre las cepas de Zapotitlán Salinas, en contraste a la secuencia de *Halomonas* del GenBank con número de acceso AB196591, con la cual se detectaron hasta 44 cambios se construyó una matriz del índice de polimorfismo entre cepas para la región parcial del gen *ectC* (225 pb) obtenida (Cuadro 2).

Cuadro 2. Índice de polimorfismo (%) entre cada cepa y la secuencia del gen *ectC* del GenBank.

	la1	I-2	la3	I-6	IVc9	Vb4	AB196591
la1							
I-2	2.66						
la3	3.11	0.44					
I-6	4.44	1.77	2.66				
IVc9	1.33	2.22	2.66	4			
Vb4	0.88	1.77	2.22	4	0.88		
AB196591	16	16.88	17.77	19.55	16	16.44	

Las secuencias presentan cambios a nivel nucleotídico, no obstante, estos no modifican en su mayoría al aminoácido a traducir ya que con el uso de codón de bacterias, los aminoácidos correspondientes presentan una hidrofobicidad similar (Cuadro 3a y 3b).

Cuadro 3a. Cambios de codón entre las secuencias nucleotídicas de las cepas de Zapotitlán Salinas, las filas sin color marcan un cambio de codón por el mismo aminoácido (sinónimo) el color amarillo marca un cambio por un aminoácido con una hidrofobicidad similar, y el color naranja marca una hidrofobicidad contraria.

Cepas	Posición	Codón	Aminoácido
Ia1/I-2	176-178	TGG/ACC	W/T
Ia1/I-2	179-181	CCT/TAT	P/Y
Ia1/I-2	215-217	CAT/CAC	H/H
Ia1/Ia3	146-148	GAG/GGG	E/G
Ia1/Ia3	176-178	TGG/ACC	W/T
Ia1/Ia3	179-181	CCT/TAT	P/Y
Ia1/Ia3	215-217	CAT/CAC	H/H
Ia1/I-6	134-136	GAA/GAT	E/D
Ia1/I-6	140-142	GAA/AAA	E/K
Ia1/I-6	143-145	GGT/AGT	G/S
Ia1/I-6	176-178	TGG/ACC	W/T
Ia1/I-6	182-184	AAT/ATT	N/I
Ia1/I-6	209-211	GAC/ACC	D/N
Ia1/I-6	215-217	CAT/CAC	H/H
Ia1/IVc9	176-178	TGG/TGC	W/C
Ia1/IVc9	185-187	AAG/AAA	K/K
Ia1/IVc9	215-217	CAT/CAC	H/H
Ia1/Vb4	176-178	TGG/TCG	W/S
Ia1/Vb4	215-217	CAT/CAC	H/H
I-2/Ia3	146-148	GAG/GGG	G/G
I-2/I-6	134-136	GAA/GAT	E/D
I-2/I-6	140-142	GAA/AAA	E/K
I-2/I-6	143-145	GGT/AGT	G/S
I-2/I-6	209-211	GAC/ACC	D/N
I-2/IVc9	176-178	ACC/TGC	T/C
I-2/IVc9	179-181	TAT/CCT	Y/P
I-2/IVc9	185-187	AAG/AAA	K/K
I-2/Vb4	176-178	ACC/TCG	T/S
I-2/Vb4	179-181	TAT/CCT	Y/P

Cepas	Posición	Codón	Aminoácido
Ia3/I-6	134-136	GAA/GAT	E/D
Ia3/I-6	140-142	GAA/AAA	E/K
Ia3/I-6	143-145	GGT/AGT	G/S
Ia3/I-6	146-148	GGG/GAG	G/E
Ia3/I-6	209-211	GAC/ACC	D/N
Ia3/IVc9	146-148	GGG/GAG	G/E
Ia3/IVc9	176-178	ACC/TGC	T/C
Ia3/IVc9	179-181	TAT/CCT	Y/P
Ia3/IVc9	185-187	AAG/AAA	K/K
Ia3/Vb4	146-148	GGG/GAG	G/E
Ia3/Vb4	176-178	ACC/TCG	T/S
Ia3/Vb4	179-181	TAT/CCT	Y/P
I-6/IVc9	134-136	GAT/GAA	D/E
I-6/IVc9	140-142	AAA/GAA	K/E
I-6/IVc9	143-145	AGT/GGT	S/G
I-6/IVc9	176-178	ACC/TGC	T/C
I-6/IVc9	182-184	ATT/AAT	I/N
I-6/IVc9	185-187	AAG/AAA	K/K
I-6/IVc9	209-211	ACC/GAC	N/D
I-6/Vb4	134-136	GAT/GAA	D/E
I-6/Vb4	140-142	AAA/GAA	K/E
I-6/Vb4	143-145	AGT/GGT	S/G
I-6/Vb4	176-178	ACC/TCG	T/S
I-6/Vb4	182-184	ATT/AAT	I/N
I-6/Vb4	209-211	ACC/GAC	N/D
IVc9/Vb4	176-178	TGC/TCG	C/S

Cuadro 3b. Cambios de codón entre las secuencias nucleotídicas de las cepas de Zapotitlán Salinas y la secuencia en la base de datos del GenBank con número de acceso AB1196591, las filas sin color marcan un cambio de codón por el mismo aminoácido (sinónimo) el color amarillo marca un cambio por un aminoácido con una hidrofobicidad similar, y el color naranja marca una hidrofobicidad contraria.

Cepas	Posición	Codón	Aminoácido	Cepa que presenta la misma mutación
Ia1/AB196591	5-7	AAT/AAC	N/N	I-2, Ia3, I-6, IVc9, Vb4
Ia1/AB196591	11-13	GAT/GAC	D/D	I-2, Ia3, I-6, IVc9, Vb4
Ia1/AB196591	23-25	TTG/CTG	L/L	I-2, Ia3, I-6, IVc9, Vb4
Ia1/AB196591	26-28	GTC/GTG	V/V	I-2, Ia3, I-6, IVc9, Vb4
Ia1/AB196591	32-34	GCT/GCC	A/A	I-2, Ia3, I-6, IVc9, Vb4
Ia1/AB196591	44/46	GGC/GGA	G/G	I-2, Ia3, I-6, IVc9, Vb4
Ia1/AB196591	47-49	AAC/CAG	N/Q	I-2, Ia3, I-6, IVc9, Vb4
Ia1/AB196591	50-52	TTC/TTT	F/F	I-2, Ia3, I-6, IVc9, Vb4
Ia1/AB196591	53-55	TCG/TCT	S/S	I-2, Ia3, I-6, IVc9, Vb4
Ia1/AB196591	56-58	TTC/TTT	F/F	I-2, Ia3, I-6, IVc9, Vb4
Ia1/AB196591	59-61	CAC/CAT	H/H	I-2, Ia3, I-6, IVc9, Vb4
Ia1/AB196591	62-64	ATT/ATC	I/I	I-2, Ia3, I-6, IVc9, Vb4
Ia1/AB196591	65-67	ACC/ACT	T/T	I-2, Ia3, I-6, IVc9, Vb4
Ia1/AB196591	68-70	CGT/CGA	R/R	I-2, Ia3, I-6, IVc9, Vb4
Ia1/AB196591	71-73	ATC/ATT	I/I	I-2, Ia3, I-6, IVc9, Vb4
Ia1/AB196591	92-94	CAC/CAT	H/H	I-2, Ia3, I-6, IVc9, Vb4
Ia1/AB196591	98-100	CAC/CAT	H/H	I-2, Ia3, I-6, IVc9, Vb4
Ia1/AB196591	104-106	AAG/AAA	K/K	I-2, Ia3, I-6, IVc9, Vb4
Ia1/AB196591	113-115	TAC/TAT	Y/Y	I-2, Ia3, I-6, IVc9, Vb4
Ia1/AB196591	143-145	GGT/GGC	G/G	I-2, Ia3, I-6, IVc9, Vb4
Ia1/AB196591	149-151	GTG/GTA	V/V	I-6, IVc9, Vb4
Ia1/AB196591	152-154	GAG/GAA	E/E	I-2, Ia3, I-6, IVc9, Vb4
Ia1/AB196591	155-157	ACC/ACA	T/T	I-2, Ia3, I-6, IVc9, Vb4
Ia1/AB196591	158-160	CTG/CTC	L/L	I-2, Ia3, I-6, IVc9, Vb4
Ia1/AB196591	161-163	GCC/GCT	A/A	I-2, Ia3, I-6, IVc9, Vb4
Ia1/AB196591	170-172	AAA/AAG	K/K	I-2, Ia3, I-6, IVc9, Vb4
Ia1/AB196591	182-184	AAT/ATC	N/I	IVc9, Vb4
Ia1/AB196591	185-187	AAG/AAA	K/K	I-2, Ia3, I-6
Ia1/AB196591	188-190	CCA/CCG	P/P	I-2, Ia3, I-6, IVc9, Vb4
Ia1/AB196591	191-193	GGC/GGT	G/G	I-2, Ia3, I-6, IVc9, Vb4
Ia1/AB196591	194-196	GAC/GAT	D/D	I-2, Ia3, I-6, IVc9, Vb4
Ia1/AB196591	200-202	TAC/TAT	Y/Y	I-2, Ia3, I-6, IVc9, Vb4
Ia1/AB196591	203-205	ATT/ATC	I/I	I-2, Ia3, I-6, IVc9, Vb4
Ia1/AB196591	206-208	CTG/CTC	L/L	Ia3, I-6, IVc9, Vb4
I-2/AB196591	176-178	ACC/TGG	T/W	Ia3, I-6
I-2/AB196591	179-181	TAT/CCT	Y/P	Ia3
I-2/AB196591	215-217	CAC/CAT	H/H	Ia3, I-6, IVc9, Vb4
Ia3/AB196591	146-148	GGG/GAG	G/E	
I-6/AB196591	134-136	GAT/GAA	D/E	
I-6/AB196591	140-142	AAA/GAA	K/E	
I-6/AB196591	143-145	AGT/GGT	S/G	
I-6/AB196591	182-184	ATT/ATC	I/I	
I-6/AB196591	209-211	ACC/GAC	N/D	
Vb4/AB196591	176-178	TCG/TGG	S/W	

Posteriormente, como resultado de la traducción *in silico* de las secuencias se generó una secuencia proteica consenso, para realizar un BLAST de proteínas, la búsqueda arrojó como resultado una proteína idéntica a la secuencia con número de acceso BAD12221.1, llamada ectoine synthase [*Halomonas sp.* TNB I20], con una cobertura del 100% con la secuencia consenso, un score máximo de 157 y un valor de error de $2e-47$.

Para la búsqueda de dominios conservados se utilizó la secuencia llamada ectoine synthase [*Halomonas sp.* TNB I20] con número de acceso BAD12221.1 - mencionada en el párrafo anterior - y con ayuda de la herramienta de dominios conservados del GenBank, las secuencias se alinearon en el programa Geneious 7.1.3 para su mejor observación.

Las secuencias resultantes de análisis pertenecen a la proteína ectC de diferentes especies de bacterias con la cuales presentó mayor similitud, éstas se presentan en orden de aparición con los números de acceso: Q0VMJ8_ALCBS, BAD12221.1, Q1Z754_PHOPR, Q2SLV8, Q2BIB4, Q45L79, Q9ZEU6, Q4U321, Q21LH6 y Q6PR31, mismos que pertenecen a las especies *Alcanivorax borkumensis*, *Halomonas sp.* TNB I20, *Photobacterium profundum*, *Hahella chejuensis*, *Neptuniibacter caesariensis*, *Cobetia marina*, *Chromohalobacter salexigens*, *Halomonas sp.*, *Saccharophagus degradans*, y *Virgibacillus pantothenicus*, respectivamente. Todas las secuencias proteicas alineadas forman parte de la superfamilia de las Cupinas. En la Figura 5, se observan los aminoácidos alineados, resaltados por su polaridad.

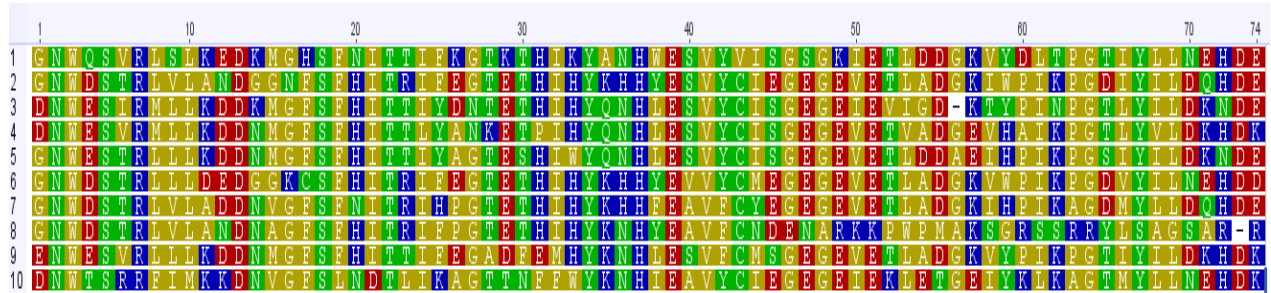


Figura 5. Dominios conservados, la secuencia 2 (*Halomonas sp.* TNB I20) es la secuencia idéntica a la consenso obtenida en el programa Geneious 7.1.3. Los colores indican la polaridad de los aminoácidos: amarillo no polar, azul polar con carga positiva, verde polar sin carga y rojo polar con carga negativa.

Para obtener el modelo teórico de la proteína se tradujo la secuencia nucleotídica de las secuencias de las cepas de Zapotitlán y se compararon en el programa Geneious 7.1.3 con las secuencias más parecidas, que fueron (*Halomonas sp.* QHL1 ectoine synthase (ectC) gene, complete cds, JX312792.1 y *Halomonas sp.* NT N60 ectC gene for ectoine synthase partial cds, strain: NT N60, AB196591.1) para obtener una secuencia consenso siguiendo el método descrito por Moya 2003 (Cuadro 4).

Cuadro 4. Modelos teóricos que puede presentar la proteína ectoína sintetasa, ordenados con respecto al valor QMEAN.

Modelo	Identidad de la secuencia	Estado del Oligo	Encontrado por	Método	Resolución	Similitud de la secuencia	Convertura	Descripción	QMEAN4
3i2h.1.D	26.15	homo-tetramero	HHblits	Rayos X	1.85Å	0.34	0.88	Azúcar Fosfato isomerasa putativa	-1.96
2h0v.1.A	28.79	homo-dimero	HHblits	Rayos X	2.60Å	0.34	0.89	Quercitina 2,3-dioxigenasa	-1.97
4rd7.1.A	24.29	homo-dimero	HHblits	Rayos X	1.57Å	0.31	0.95	Cupina 2 con un dominio conservado de barril	-2.21

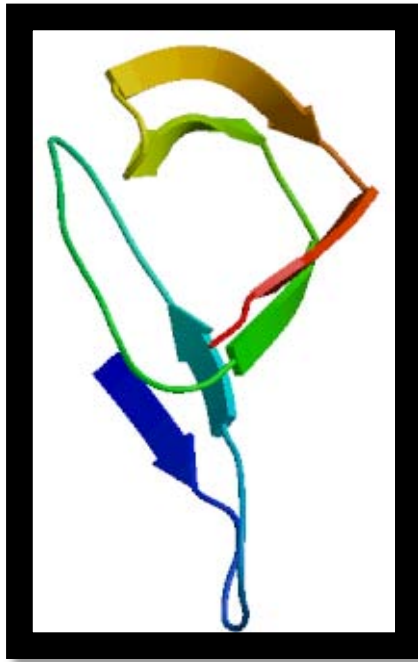


Figura 6. Azúcar fosfato isomerasa putativa

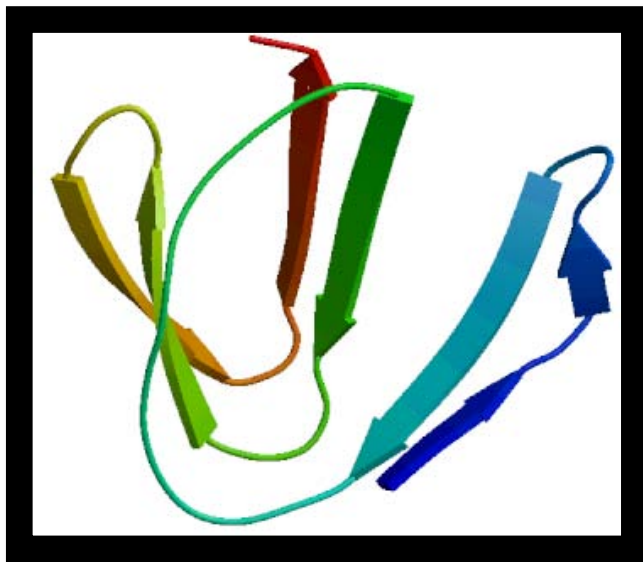


Figura 7. Quercitina 2,3 dioxigenasa

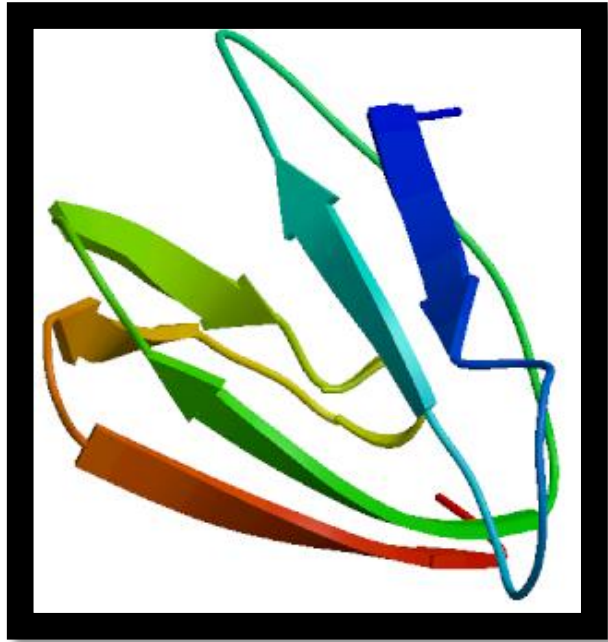


Figura 8. Cupina 2 proteína con un dominio conservado de barril

Los Modelos de la Figuras 6: Azúcar fosfato isomerasa putativa de *Acidithiobacillus ferrooxidans* (JCSG, 2014), Figura 7: Quercitina 2,3 dioxigenasa de *Bacillus subtilis* (JCSG, por publicar) y Figura 8: Cupina 2 proteína con un dominio de barril conservado de *Salinispora arenicola* CNS-205, se encuentran dentro de la superfamilia de las Cupinas.

DISCUSIÓN

Para la obtención de la secuencia parcial del gen *ectC* de cepas de las *Halomonas* de Zapotitlán, se necesitó de un proceso de amplificación en dos partes; en primera instancia, se observaron de 5 a 7 bandas (Figura 2). Para las cepas bacterianas de este estudio, se utilizaron primers que están degenerados en cuatro posiciones en la secuencia Directa y 6 en la secuencia del Reverso, por lo que esta inespecificidad pudo permitir la amplificación de regiones que no corresponden al gen *ectC*, aunque al estar diseñados sobre regiones conservadas reconocen otras que presentan algún grado de similitud en las región correspondientes a éstos primers, tal es el caso del amplicón de 800pb, el análisis de su secuencia arrojó que pertenecía a una proteína de unión a monosacáridos (datos no mostrados).

Para obtener una sola banda se requirió de una reamplificación de la banda de 300 pb (Figura 3), misma que corresponde al tamaño esperado para el amplicón de *ectC* de *Halomonas* reportado por Zanjirband *et al.*, en el 2008. Las secuencias de los amplicones obtenidos de las cepas de Zapotitlán Salinas, Puebla, al ser utilizadas en la base de datos del GenBank mostraron que pertenecían al gen *ectC* de *Halomonas* (Cuadro 1). Cabe mencionar que hasta el momento en el GenBank se han dado de alta 3158 secuencias del gen *ectC* de las cuales sólo 53 pertenecen al género *Halomonas*, lo cual representa un 1.67% del total y ninguna de estas secuencias corresponden a especies o cepas mexicanas.

La cepa Ia1, Ia3, I-6, IVc9 y Vb4 tuvieron mayor homología con *Halomonas sp.* QHL1 ectoine synthase (*ectC*) gene, complete cds, con

número de acceso JX312792.1, teniendo una similitud del 87, 85, 79, 85 y 84% respectivamente. La cepa *Halomonas* sp. QHL1 aislada del lago Qinghai, en China y fue utilizada para clonar el operón *ectABC* y transformar a *E. coli*, y obtener una cepa transformada resistente a altas concentraciones de sal, no obstante, esta cepa producía menor cantidad de ectoína que la cepa de *Halomonas* sp. QHL1. (Zhu *et al.*, 2014).

La cepa I-2 tuvo mayor homología con *Halomonas* sp. NT N60 *ectC* gene for ectoine synthase partial cds, strain: NT N60, con número de acceso AB196591.1 y una similitud del 83%. (Okamoto y Naganuma, dato sin publicar, tomado del GenBank).

En el Cuadro 2, se muestra el índice de polimorfismo entre las cepas de Zapotitlán Salinas, donde la cepa I-6 es la más diferentes entre ellas con un 4.44% y las que presentan mayor similitud son I-2 y I-3 con tan sólo 0.44%. Con respecto a la secuencia con número de acceso AB196591 que es la más parecida a estas secuencias se observa que tienen un índice de polimorfismo mayor a 16%. Esto indica que tiene una distancia evolutiva mayor, lo que pudiera estar relacionado al lugar de procedencia de la muestra, ya que las cepas utilizadas en este trabajo se aislaron de la parte continental en México y la secuencia de la base de datos fue aislada del sedimento del mar en Japón.

En la posición 74-94 no existen cambios en ninguno de los nucleótidos, conservándose en todas las secuencias, lo que puede ser el sitio catalítico de la enzima o algún tipo de sitio de reconocimiento. Los genes involucrados en la síntesis de ectoína están altamente conservados en diversas especies de bacterias que residen en ambientes donde hay

cambios de salinidad y se observa con frecuencia que la biosíntesis de ectoína es fuertemente inducida en altas concentraciones de sal (Costa *et al.*, 1998), tal es el caso del Valle de Zapotitlán Salinas, donde los sedimentos y agua de las salineras del sitio de colecta las cepas presentan una salinidad entre 16 a 40 UPS (Unidades prácticas de Salinidad) (16 a 40 g/L de sal), esta concentración de sales permite el crecimiento de cepas de *Halomonas* que integran en su genoma a los genes *ectC*.

Se observa en el Cuadro 3a, diez de los cambios de codón entre las cepas dan el mismo aminoácido, 40 cambios son aminoácidos de la misma naturaleza de hidrofobicidad y sólo 5 son cambios por aminoácidos de hidrofobicidad contraria. Mientras en el Cuadro 3b se encontraron hasta 34 cambios de codón que con respecto a la cepa del GenBank traducen el mismo aminoácido, 9 cambios son aminoácidos de la misma naturaleza de hidrofobicidad y sólo un cambio por un aminoácido de hidrofobicidad contraria G/E, la glicina aminoácido hidrofóbico y el ácido glutámico es hidrofílico. La naturaleza de los cambios presentados a nivel de aminoácidos sugiere que no se afecta a la proteína y mantiene su función, ya que las cepas crecen en sitios que reportan concentraciones hasta 6 UPS (de NaCl 6g/L) en las salineras de Zapotitlán Salinas (Novelo, 2011) de donde fueron aisladas las cepas.

En cuanto a identificación de dominios conservados en los genes *ectC*, se hizo un primer acercamiento buscando el dominio conservado del mismo directamente en la herramienta de dominios conservados del GenBank, lo cual mostró que *ectC* existe en muchas especies de bacterias, entre ellas *Vibrio cholerae* (Heidelberg *et al.*, 2000) especie de

importancia médica, *Alcanivorax borkumensis* (Schneike *et al.*, 2006) bacteria que se reporta como capaz de degradar petróleo, *Neptuniibacter caesariensis* bacteria es de origen marino, *Hahella chejuensis* (Jeong *et al.*, 2005) usada como alguicida.

Como se ha establecido éste gen se encuentra distribuido en una gran variedad de bacterias sometidas a condiciones de estrés salino y algunas de ellas poseen potencial biotecnológico, como lo demuestra el número de patentes de los diferentes productos lanzados al mercado por Osmoshield® (2014) donde se utilizan dos osmoprotectores: la Ectoína y Trehalosa para protección de la estructura nativa de las proteínas celulares, favoreciendo el equilibrio osmótico y la vitalidad de las células. Brill Pharma (2015) lanzó al mercado 'Hylo-Dual', una combinación de ectoína y ácido hialurónico cuyo objetivo es prevenir y reducir los síntomas oculares de la alergia, puesto que la ectoína forma una barrera fisiológica en la superficie ocular que bloquea a los alérgenos, protegiendo a la córnea de las irritaciones que estos le pueden causar. *Bitop. Extremolytes for life. Ectoin® Rhinitis Nasal Spray –A cold* utilizada en gotas para los ojos y sprays nasal Ectoin® cuida la humedad de los epitelios, Ectoin® y Ectoin® Hydro forman un complejo con el agua que se encuentra fuera de las células manteniendo así la humedad.

En la Figura 5 se observa que existen regiones altamente conservadas, ya que la mayoría de los aminoácidos se mantiene o mantiene una naturaleza de polaridad similar, solamente algunas regiones se ve un cambio por un aminoácido de polaridad distinta, conservando la naturaleza de la proteína. Pese a los diversos trabajos consultados no se

logró documentar la estructura terciaria por lo que se recurrió a la herramienta de modelado por computadora con base en la secuencia de aminoácidos.

Como se muestra en el Cuadro 4 son tres los modelos que más coinciden con la secuencia de aminoácidos obtenida en la traducción de las secuencias de nucleótidos de las *Halomonas*. Los modelos teóricos de la secuencia corresponden a la familia de las cupinas del latín “cupa” pequeño barril (Dunwell, 1998).

El primer modelo Figura 6 Manosa - 6 -fosfato isomerasa o fosfomanosa isomerasa (PMI) es la enzima que cataliza la interconversión de manosa - 6 -fosfato y fructosa - 6 -fosfato. En eucariotas PMI está implicado en la síntesis de GDP - manosa, un constituyente de N- y O- glicanos ligados y anclajes GPI, en procariontes participa en una variedad de vías, incluyendo la biosíntesis de polisacárido capsular y el metabolismo de D-manosa. PMI pertenece a la superfamilia de las Cupinas cuyas funciones varían desde actividades isomerasa y epimerasa que participan en la modificación de hidratos de carbono de la pared celular en bacterias y plantas, a las proteínas no enzimáticas de almacenamiento en semillas de plantas (Dunwell *et al.*, 2004).

El segundo modelo Figura 7 pertenece a la familia parecida a Quercetina 2,3 dioxigenasa (2,3 QD) es una dioxigenasa que contiene en el centro Cu^{+2} , tiene como función la inserción de oxígeno molecular en la quercetina en la ruta de la degradación de flavonoles polifenólicos, esta reacción catalizada por la glicoproteína 2,3 QD es única, en ella se involucra un centro de Cu^{+2} en el catabolismo, ya que normalmente la

incorporación de O₂ a un sustrato biológico se lleva a cabo mediante enzimas que utilizan hierro como cofactor. (Fusetti *et al.*, 2002). Comparando la secuencia consenso obtenida y la secuencia de esta proteína se encontraron similitudes por ejemplo en la posición 31-33 existe un dominio H I H que forma una hélice β , posición 39 y 40 residuos Y E junto con residuos Y C posición 42 y 43 forman parte de otra hélice β y posición 64 y 65 residuos G D comienzan otra hélice β .

El tercer modelo Figura 8 conserva un dominio de barril como su nombre lo dice, pertenece al dominio de las cupinas del latín “cupa” pequeño barril (Dunwell, 1998). Los modelos anteriores presentan una semejanza en la estructura de la proteína, entre algunas de las funciones que se le atribuyen a proteínas similares se involucran la respuesta al estrés en organismos tanto en procariontes como eucariontes, otra función es impedir la pérdida de agua o desecación en plantas superiores dando una estabilidad en altas temperaturas, lo cual es ejemplificado en las proteínas de almacenamiento en la semilla, por un grupo muy conocido de proteínas no enzimáticas, las cuales contienen en su mayoría un solo residuo conservado His y sin actividad enzimática conocida (Bäumelin *et al.*, 1995) y resistencia a las proteasas (Uberto y Moomaw 2013), lo cual muestra una función análoga con la ectoína, esto infiere que la superfamilia de las cupinas, la cual se caracteriza por conservar una estructura bien definida en todas las proteínas (Dunwell *et al.* 2000), confiere una función análoga a las proteínas de los organismos, lo que concuerda con Uberto y Moomaw 2013 quienes concluyen que basados en el agrupamiento de estructura y secuencia de los miembros de la familia de las Cupinas refleja un agrupamiento funcional.

CONCLUSIONES

- El gen *ectC* está presente en las seis cepas de *Halomonas* sp. de Zapotitlán Salinas, por lo que este es el primer reporte de la secuencia nucleotídica parcial de la ectoína sintasa en México.
- Las cepas de *Halomonas* aisladas del Valle de Zapotitlán están estrechamente relacionadas con las secuencias reportadas para este género. Se conserva en particular 20 nucleótidos que son idénticos entre cepas e incluso con otras *Halomonas* depositadas en el GenBank, lo que sugiere que es una región importante para la función.
- Dentro de las 6 cepas analizadas se detectaron mutaciones en las secuencias del gen *ectC* de hasta un 19.5% (44 cambios) con respecto a la secuencia de *Halomonas* sp. NT60 (GenBank AB196591). No obstante, el análisis bioinformático reveló que el cambio de base en el triplete no modifica el aminoácido traducido, así, en lo general la secuencia de la enzima ectoína sintasa se mantiene en las cepas de *Halomonas* de Zapotitlán Salinas.
- Los dominios conservados de EctC de las cepas de las *Halomonas* estudiadas, la agrupan dentro de la superfamilia de las cupinas.
- Los tres modelos teóricos de proteínas a los que se ajustan las secuencias traducidas *in silico* de las cepas estudiadas, consideradas bajo el criterio del valor QMEAN más cercano a cero, señala que pertenecen a la superfamilia de las Cupinas, éstas proteínas tienen una estructura definida tipo barril y participan en respuesta al estrés por altas temperaturas o pérdida

de agua, lo que sugiere que esta estructura mantiene una función en los diferentes organismos, desde bacterias hasta plantas. Así, la EctC de *Halomonas* de Zapotitlán Salinas puede estar involucrada en la sobrevivencia de estas bacterias en tales condiciones de salinidad.

LITERATURA CITADA

Arahal, D.R., Ludwig, W., Schleifer, K.H., & Ventosa, A. (2002). Phylogeny of the family *Halmomonadaceae* based on 23S and 16S rDNA sequence analyses. *Int. J. Syst. Microbiol.* 52: 241-249.

Arahal, D.R., & Ventosa, A. (2006). The family Halomonadaceae. En: Dworkin, M., Falkow, S., Rosenberg, E., Schleifer, K.H., & Stackebrandt, E. (eds). *The Prokaryotes*, 3rd edition, vol. 6 (Proteobacteria: Gamma Subclass). Springer, New York, 811-835.

Arahal, D.R., Vreeland, R.H., Litchfield, C.D., Mormile, M.R., Tindall, B.J., Oren, A., Béjar, V., Quesada, E., & Ventosa, A. (2007). Recommended minimal standards for describing new taxa of the family Halomonadaceae. *Int J Syst Evol. Microbiol.* 57: 2436-2446.

Benkert, P., Tosatto, S.C.E., & Schomburg, D. (2008). QMEAN. A comprehensive scoring function for model quality assessment. *Prot.* 71:1:261-277.

Brown, A.D. 1976. Microbial water stress. *Bacteriol. Rev.* 40: 803-846.

CONABIO. (2013). RTP-121, Valle de Tehuacan-Cuicatlan. Disponible en http://www.conabio.gob.mx/conocimiento/regionalizacion/doctos/rtp_121.pdf vi: 13/08/13.

da Costa, M.S., Santos, H., & Galinski, E.A. (1998). An overview of the role and diversity of compatible solutes in Bacteria and Archea. *Adv. Biochem. Eng. Biotech.* 61: 117-153.

Dunwell, J.M. (1998). Cupins: a new superfamily of functionally-diverse proteins that include germins and plant seed storage proteins. *Biotech. Genet Engin. Rev.* 15: 1-12.

Dunwell, J.M., Purvis, A., & Khuri, S. (2004). Cupins: the most functionally diverse protein superfamily? *Phytochem.* 65: 7-17.

Dunwell, J.M., Khuri, S., & Gane, P.J. (2000). Microbial relatives of the seed storage proteins of higher plants: Conservation of structure and diversification of function during evolution of the cupin superfamily. *Microbiol. Mol. Biol. Rev.* 64 (1): 153-179.

Fusetti, F., Schröter, K.H., Steiner, R.A., van Noort, P.I., Pijning, T., Rozeboom, H.J., Kalk, K.H., Egmond, M.R., & Dijkstra, B.W. (2002). Crystal structure of the Copper-Containing Quercetin 2,3-Dioxygenase from *Aspergillus japonicus*. *Structure.* 10: 259-268.

Galinski, E.A., Pfeiffer, H.P., & Trüper, H.G. (1985). 1,4,5,6-Tetrahydro-2-methyl-4-pyrimidinecarboxylic acid. *Eur J Biochem* 149: 135-139.

Garrity G.M., Bell, J.A., & Lilburn, T. (2005). Class III. Gammaproteobacteria class. nov. In: Brenner, D.J., Krieg, N.R., Staley, J.T., & Garrity, G.M. (eds.). *Bergey's Manual of Systematic Bacteriology*, 2nd edition, vol. 2 (The Proteobacteria), part B (The Gammaproteobacteria). Springer, New York, 1.

Heidelberg, J.F., Eisen, J.A., Nelson, W.C., Clayton, R.A., Gwinn, M.L., Dodson, R.J., Haft, D.H., Hickey, E.K., Peterson, J.D., Umayam, L., Gill, S.R., Nelson, K.E., Read, T.D., Tettelin, H., Richardson, D., Ermolaeva, M.D., Vamathevan, J., Bass, S., Qin, H., Dragoi, I., Sellers, P., McDonald, L., Utterback, T., Fleishmann, R.D., Nierman, W.C., White, O., Salzberg, S.L., Smith, H.O., Colwell, R.R., Mekalanos, J.J., Venter, J.C., & Fraser, C.M. (2000). DNA sequence of both chromosomes of the cholera pathogen *Vibrio cholera*. *Nat.* 406: 477- 483.

Jeong, H., Yim, J.H., Lee, D., Choi, S.H., Park, Y.K., Yoon, S.H., Hur, C.G., Kang, H.Y., Kim, D., Lee, H.H., Park, K.H., Park, S.H., Park, H.S., Lee, H.K., Oh, T.K., & Kim, J.F. (2005). Genomic blueprint of *Hahella chejuensis*, a marine microbe producing an algicidal agent. *Nucleic. Acids. Res.* 33(22): 7066-7073.

Joint Center for Structural Genomics (JCSG). (To be published). Crystal Structure of (2636545) from *Bacillus Subtilis* at 2.60 Å Resolution.

Joint Center for Structural Genomics (JCSG). (2014). Crystal structure of Putative sugar phosphate isomerase (AFE 0303) from *Acidithiobacillus ferrooxidans* ATCC 23270 at 1.85 Å resolution. wwPDB X-ray Structure Validation Summary Report. 16 pp.

Lapidot, A., Ben-Asher, E., & Eisenstein, M. (1995). Tetrahydropyrimidine derivatives inhibit binding of a Tat-like, arginine-containing peptide, to HIV TAR RNA in vitro. *FEB Lett.* 367: 33–38.

Litchfield, C.D. 1998. Survival strategies for microorganisms in hypersaline environments and their relevance to life on early Mars. *Met Plan. Science*. 33(4): 813-819.

Llamas, I., Béjar, V., Martínez-Checa, F., Martínez-Cánovas, M.J., Molina, I., & Quesada, E. (2011). *Halomonas stenophila* sp. nov., a halophilic bacterium that produces sulphate exopolysaccharides with biological activity. *Int J Syst Evol Microbiol*. 61: 2508-2514.

Lo, C.C., Bonner, C.A., Xie, G., D'Souza, M., & Jensen, R.A. (2009). Cohesion group approach for evolutionary analysis of aspartokinase, an enzyme that feeds a branched network of many biochemical pathways. *Microbiol Mol Biol Rev*. 73(4): 594-651.

Marchler, B.A., Zheng, C., Chitsaz, F., Derbyshire, M.K., Geer, L.Y., Geer, R.C., Gonzales, N.R., Gwadz, M., Hurwitz, D.I., Lanczycki, C.J., Lu, F., Lu, S., Marchler, G.H., Song, J.S., Thanki, N., Yamashita, R.A., Zhang, D., & Bryant, S.H. (2013). CDD: conserved domains and protein three-dimensional structure. *Nucleic Acids Res*. 41: 348-352.

Margesin, R., & Schinner, F. (2001). Potential of halotolerant and halophilic microorganisms for biotechnology. *Extremophiles*. 5: 73-83.

Morbach, S., & Krämer, R. (2002). Body Shaping under Water Stress: Osmosensing and Osmoregulation of Solute Transport in Bacteria. *Chem BioChem*, 3: 384-397.

Mata, J.A., Martínez-Cánovas, M.J., Quesada, E., & Béjar, V. (2002). A detailed phenotypic characterization of the type strains of *Halomonas* species. *Syst Appl Microbiol* 25:360-375.

Min-Yu, L., Ono, H., & Takano, M. (1993). Gene cloning of ectoine synthase from *Halomonas* sp. *Annu Rep Int Centre Cooperative Res Biotechnol Japan*. 16: 193-200.

Moya, G.A.A. (2003). Predicción de estructura de proteínas empleando software libre. *Encuentros Biol*. 87.

Novelo, E. (2011). Cyanoprokaryota en Flora del Valle de Tehuacán. Instituto de Biología. UNAM. 90: 1- 96

Ono, H., Sawada, K., Khunajakr, N., Tao, T., Yamamoto, M., Hiramoto, M., Shinmyo, A., Takano, M., & Murooka, Y. (1999). Characterization of biosynthetic enzymes for ectoine as a compatible solute in a moderately halophilic eubacterium, *Halomonas elongata*. *J Bacteriol*. 181: 91-99.

Owen, R.J., & Pitcher, D. (1985). Current methods for estimating DNA base composition and levels of DNA-DNA hybridization. *In Chemical Methods in Bacterial Systematics*. 67–93 pp. Edited by M. Goodfellow & E. Minnikin. London: Academic Press.

Pastor, J.M., Salvador, M., Argandoña, M., Bernal, V., Reina-Bueno, M., Csonka, L.N., Iborra, J.L., Vargas, C., Nieto, J.J., & Cánovas, M.(2010). Ectoines in cell stress protection: Uses and biotechnological production. *Biotech Adv*. 28: 782-801.

Peters, P., Galinski, E.A., & Trüper, H.G. (1990). The biosynthesis of ectoine. *FEMS Microbiol Lett*. 71: 157-162.

Quesada, E., Ventosa, A., Ruiz-Berraquero, F., & Ramos-Cormenzana, A. (1984). *Deleanda halophila*, a new species of moderately halophilic bacteria. *Int J Syst Bacteriol* 34: 287-292.

Ramírez, N., Sandoval, A.H., y Serrano, J.A. (2004). Las bacterias halófilas y sus aplicaciones biotecnológicas. *Rev. Soc. Ven. Microbiol.* 24:2: 12-23.

Reuter, K., Pittelkow, M., Bursy, J., Heine, A., Craan, T., & Bremer, E. (2010). Synthesis of 5-Hydroxyectoine from Ectoine: Crystal Structure of the Non-Heme Iron(II) and 2-Oxoglutarate-Dependent Dioxxygenase EctD. *PLoS ONE*. 5(5): e10647.

Roberts, M.F. (2005). Organic compatible solutes of halotolerant and halophilic microorganisms. *Salin Syst.* 1(5): 1-30.

Rodicio, M.R. y Mendoza, M.C. (2004). Identificación bacteriana mediante secuenciación del ARNr 16S: fundamento, metodología y aplicaciones en microbiología clínica. *Enferm Infecc Microbiol Clin.* 22(4): 238-245.

Rosey, E.L., & Stewart, G.C. (1992). Nucleotide and deduced amino acid sequences of the *lacR*, *lacABCD*, and *lacFE* genes encoding the repressor, tagatose 6-phosphate gene cluster, and sugar-specific phosphotransferase system components of the lactose operon of *Streptococcus mutans*. *Jour. Bacteriol.* 174:19:6159-6170.

Schneiker, S., Martins dos Santos, V.A., Bartels, D., Bekel, T., Brecht, M., Buhrmester, J., Chernikova, T.N., Denaro, R., Ferrer, M., Gertler, C., Goesmann, A., Golyshina, O.V., Kaminski, F., Khachane, A.N.,

Lang, S., Linke, B., McHardy, A.C., Meyer, F., Nechitaylo, T., Puhler, A., Regenhardt, D., Rupp, O., Sabirova, J.S., Selbitschka, W., Yakimov, M.M., Timmis, K.N., Vorholter, F.J., Weidner, S., Kaiser, O., & Golyshin, P.N. (2006). Genome sequence of the ubiquitous hydrocarbon-degrading marine bacterium *Alcanivorax borkumensis*. *Nat. Biotech.* 24(8): 997-1004.

Schubert, T., Maskow, T., Benndorf, D., Harms, H., & Breuer, U. (2007). Continuous synthesis and excretion of the compatible solute ectoine by a transgenic nonhalophilic bacterium. *Appl. Environ. Microbiol.* 73(10): 3343-3347.

Tan, K., Gu, M., Clancy, S., Phillips Jr, G.N., Joachimiak, A. (To be published). The crystal structure of a Cupin 2 conserved barrel domain protein from *Salinispora arenicola* CNS-205.

Uberto, R., & Moomaw, E.W. (2013). Protein Similarity Networks Reveal Relationships among Sequence, Structure, and Function within the Cupin Superfamily. *Plos One.* 8(9): e74477

Van den Burg, B. (2003). Extremophiles as a source for novel enzymes. *Curr Opin Microbiol.* 6: 1-6.

Wood, J.M., Bremer, E., Csonka, L.N., Krämer, R., Poolman, B., Van der Heide, T., & Smith, L.T. (2001). Osmosensing and osmoregulatory compatible solutes accumulation by bacteria. *Comp Biochem Physiol.* 130: 437-60.

Zanjirband, M., Golban, N., & Kermanshahi, R.K. (2008). Detection of the *ectC* gene in *Halomonas* strains by polymerase chain reaction. *Iran J of Biotech.* 6(3): 181-185.

Zhu, D., Liu, J., Han, R., Shen, G., Long, Q., Wei, X., & Liu, D. (2014). Identification and characterization of ectoine biosynthesis gene and heterologous expression of the *ectABC* gene cluster from *Halomonas* *sp.* QHL1, a moderately halophilic bacterium isolated from Qinghai lake. *J of Microbiol.* 53(2): 139-147.

ANEXO 1

Nomenclatura de cepas

Sitio de colecta	Numero de cepa de cada sitio
I	a1 , 2, a3, 6
II	
III	
IV	c9
V	b4

ANEXO 2

Las etapas necesarias para el modelado por homología son básicamente cuatro:

- Identificación de estructuras conocidas (moldes) relacionadas con la secuencia diana. Se emplean métodos de comparación de secuencias como FASTA, BLAST o PSI-BLAST [<http://www.ncbi.nlm.nih.gov/BLAST/>].
- Alineamiento de la diana con los moldes. Es la etapa más importante y la más delicada ya que la construcción del modelo se realizará conforme a este alineamiento. En este paso se emplean programas típicos de alineamiento de secuencias (GENEIOUS 7.1.3)
- Construcción del modelo. Existen varias aproximaciones para construir las coordenadas espaciales de la secuencia diana desde el alineamiento realizado. Tenemos ProModII en el servidor SWISS-MODEL [<http://www.expasy.org/swissmod/SWISS-MODEL.html>].
- Evaluación del modelo. La información que se puede obtener del modelo depende de su calidad, de modo que es importante poder evaluarla. Existen muchas pruebas que se pueden realizar sobre un modelo que incluyen comprobaciones estéricas, químicas, representaciones, etc.

Para la evaluación del modelo se toma en cuenta el valor Qualitative Model Energy Analysis (QMEAN por sus siglas en inglés) el cual es una calificación de la descripción del mejor aspecto geométrico de la

estructura de la proteína. Para obtener el valor se usan cinco diferentes aspectos estructurales:

-El ángulo de torsión sobre tres aminoácidos consecutivos.

-Un alineamiento potencial de nivel de residuos dependiente de la distancia específica de estructura secundaria, se utiliza para evaluar las interacciones de largo alcance.

-Un potencial de solvatación describe el acomodo de los residuos en un solvente.

-Dos términos simples que describen la concordancia de la estructura secundaria predicha y calculada y accesibilidad de disolvente, respectivamente,

Una variedad de diferentes implementaciones son investigados y varios enfoques para combinar y optimizarlos aún se discuten (Benkert *et al.*, 2008).

ANEXO 3

Aminoácido	Código de tres letras	Código de una letra
Alanina	Ala	A
Arginina	Arg	R
Asparagina	Asn	N
Ácido aspártico	Asp	D
Cisteína	Cys	C
Glutamina	Gln	Q
Acido glutámico	Glu	E
Glicina	Gly	G
Histidina	His	H
Isoleucina	Ile	I
Leucina	Leu	L
Lisina	Lys	K
Metionina	Met	M
Fenilalanina	Phe	F
Prolina	Pro	P
Serina	Ser	S
Treonina	Thr	T
Triptófano	Trp	W
Tirosina	Tyr	Y
Valina	Val	V