Comité Editorial: González Andrade, M.; Hernández Alcántara G.; Martínez González, J.J.; Ramírez Silva, L. H. y Vilchis Landeros, M.M.
© ISSN-0188-137X

Versión electrónica en http://bq.facmed.unam.mx/tab/

MENSAJE BIOQUÍMICO

Mens. Bioquim. 44 (2020) 87-95

Memoria del 47º Taller de Actualización Bioquímica, Facultad de Medicina; UNAM

Superfamilia de proteínas dehalogenasa de haloácido, la extensa radiación evolutiva de un plegamiento antiguo.

Haloacid dehalogenase protein superfamily, a wide evolutionary radiation from an ancient fold.

Rodríguez-Sotres, Rogelio[1*] y Valencia-Turcotte, Lilián Gabriela[1].

1. Departamento de Bioquímica, Facultad de Química, UNAM

*Correspondencia. Departamento de Bioquímica, Facultad de Química, UNAM, Avenida Universidad 3000, Ciudad Universitaria, Coyoacán, CDMX, CP 04510 Tel. +52 (55) 5622-5285, sotres@unam.mx

Resumen

La superfamilia de proteínas de las dehalogenasas de haloácido (HAD) es un grupo diversificado de proteínas bien conocido. El plegamiento se considera antiguo, pues parece provenir de las bacterias primigenias y todo genoma reportados contienen varios genes que codifican para miembros de esta superfamilia, algunos como un dominio de una proteína más grande. Sus funciones incluyen fosfatasas inespecíficas y específicas de metabolitos y de fosfoproteínas, pirofosfatasas, ATPasas, fosfotransferasas, epóxido hidrolasas, dehalogenasas y más. Además de un papel en el metabolismo, en animales se les ha implicado recientemente en la señalización, el ciclo celular y el cáncer. En plantas, la evidencia sugiere que continúan diversificándose, especialmente en especies agrícolas, pero sus funciones no están bien estudiadas. La fosfatasa del fosfato de metiltiopentenilo de *Bacillus subtilis* es única pues presenta un dedo de zinc en el sitio activo y a pesar de la escasa conservación de secuencia, tal característica parece presente en una proteína vegetal de función debatida, cuya estructura modelamos y dicho modelo aportó datos sobre peculiaridades en su acción catalítica. Dada la diversidad de función y secuencia, la superfamilia HAD es un paradigma interesante desde el punto de vista evolutivo y de la ingeniería de proteínas.

Abstract

The haloacid dehalogenase (HAD) superfamily of proteins is considered as the most diverse group of proteins. It is thought to be an ancient fold, already present in the first bacteria, and currently encoded by several genes in all known genomes, in some cases as a domain of a larger protein. Their functions include phosphatases of high and low specificity, acting on metabolites or phosphoproteins, pyrophosphatases, phophotransferases, and among other, epoxide and halogen hydrolases. The methyl-thiopentenyl-phosphate phosphatase form *Bacillus subtilis* has a distinctive zinc finger at the active site, and this unique feature seems to be conserved in a plant protein, despite the rather low conservation of sequence. However, the function of this last plant protein is still under debate and its structure is not known, so we produced a model and made inferences on its possible catalytic mechanism. Finally, the three-dimensional folding pattern of HAD superfamily stands out as a case of study in protein evolution and engineering, given the diversity of function and sequence it can support.

Palabras clave: Dehalogenasas de haloácido; PFAM; fosfatasas; plegamiento de proteínas; relación estructura-función de proteínas.

Key words: Haloacid dehalogenases; PFAM; phosphatases; protein folding; protein structure-function relationship.

Introducción

Con solo 20 bloques fundamentales, las proteínas pueden dar lugar a un número inmenso de combinaciones de secuencia, muchas de las cuales no han sido aún exploradas por la naturaleza. En cambio, la diversidad de plegamientos observados hasta ahora es mucho menos diverso [1]. Esto sugiere que un mismo plegamiento puede obtenerse con secuencias distintas, dando lugar así a funciones distintas para dos proteínas de igual plegamiento [2,3]. Esta característica ha sido explotada por la naturaleza en un largo proceso evolutivo, de manera que muchas proteínas modernas con funciones distintas, se originaron a partir de un mismo gene ancestral; dado lo cual, conservan características en su secuencia y en su plegamiento que permiten reconocer su origen común [3]. Mas aún, a pesar de la diversidad de secuencias que adoptan el mismo patrón de plegamiento, usando métodos estadísticos sofisticados, llamados modelos ocultos de Markov (HMM), resulta posible reconocer una firma de secuencia asociada a cada plegamiento [5]. Estos métodos se han empleado para agrupar los dominios estructurales de las proteínas en una organización jerárquica de clanes o superfamilias, cuyos miembros son a su vez familias, cada una de las cuales agrupa una multitud de secuencias relacionadas evolutivamente [6].

La superfamilia denominada "clan halogenasa de haloácido" (cl-HAD) toma su nombre de una enzima de *Pseudomonas sp*. que puede remover grupos halógeno de ácidos alcanoícos halogenados (PDB 1JUD) [7]. Pero la superfamilia cl-HAD comprende 23 familias y 248368 miembros distintos (hasta feb 2020) distribuidos en 9956 especies de seres vivos, la mayoría de los cuales presentan una actividad enzimática distinta; siendo además una de las superfamilias más extensa, entre las 643 identificadas en la base de datos PFAM [6]. La gran mayoría de estas proteínas son hidrolasas o fosfotransferasas y se caracterizan por un mecanismo catalítico que involucra un residuo de ácido aspártico que actúa como nucleófilo, parecen formar un intermediario covalente de aspartilfosfato y, además, conservan un núcleo plegamiento mínimo, que parece ser muy adaptable, dada la diversidad de funciones que soporta, con una conservación de secuencia realmente escasa.

Entre las familias miembro del cl-HAD (Tabla 1) se encuentran: ATPasa-P de calcio, como la del retículo sarcoplásmico; nucleotidasas, de varias clases y con diferencias de especificidad; pirofosfatasas, de al menos 2 tipos; fosfatasas, tanto de alta como de baja especificidad; la fosfatidilglicerolfosfato (PGP) fosfatasa, involucrada en la síntesis de cardiolipina en mitocondria; fosfohidrolasas, asociadas a la síntesis de DNA, como la polinucleótido 5'-cinasa/3'-fosfatasa (PNK3P), las proteínas Prune de *Drosophila melanogaster* y la enzima RecJ bacteriana; fosfomanomutasa, involucrada en la reparación de proteínas glicosidadas, así como dos grupos de proteínas con función desconocida (DUF).

Desafortunadamente, también hay traslape funcional entre varios grupos, lo que complica la clasificación y la nomenclatura. Es decir, frecuentemente los métodos de agrupamiento por HMM ponen en grupos separados a proteínas que tienen la misma actividad enzimática y papel biológico semejante en dos organismos diferentes. En esos casos, es difícil inferir si se trata de dos proteínas cuya evolución las llevó a converger en la misma actividad y papel biológico, siguiendo caminos diferentes (evolución convergente) o si un ancestro común mantuvo su estructura, actividad y papel fisiológico, pero sufrió mutaciones en el proceso evolutivo que llevó a la separación de las especies en que se encuentra, actualmente (radiación evolutiva, o evolución divergente). Pero además, también hace difícil poner un nombre a cada familia, por lo que los nombres en la tabla 1 dicen poco sobre las características de las proteínas en cada familia.

Razgos estructurales conservados en el clan dehalogenasa de haloácido

La estructura de las proteínas cl-HAD se caracteriza por su enorme diversidad, al grado de que la semejanza de secuencia entre dos miembros de una misma familia del clan puede ser tan baja como 15% al comparar ortólogos correspondientes a Bacteria y Eucariontes. Más aún, la semejanza de secuencia entre los ortólogos correspondientes de animales y plantas suele caer en el rango de 10-35%. Es decir, a pesar de que el PDB contiene miembros representativos de varias familias cl-HAD, el modelar la estructura de estas proteínas resulta ser un problema difícil, pues el encontrar una referencia adecuada se dificulta enormemente. El proceso solo se facilita, cuando la

© 2020 Mensaje Bioquímico. Todos los derechos reservados. ISSN-0188-137X
Comité Editorial: González Andrade, M.; Hernández Alcántara G.; Martínez González, J.J.;
Ramírez Silva, L. H. y Vilchis Landeros, M.M. *Publicado por el Departamento de Bioquímica de la Facultad de Medicina; UNAM.*

secuencia objetivo del modelado proveniente de un organismo filogenéticamente cercano al organismo cuya proteína fue resuelta experimentalmente y que servirá como referencia. Frecuentemente, esta no es la situación encontrada, pues los objetivos de modelado más interesantes son las proteínas de las que sabemos poco.

Tabla 1. Familias miembro del clan HAD en PFAM (CL0137)*.

Familia	Distribución‡	Código EC	PDB	Miembros*
Purín-5' nucleotidasa (PF05761)	An,Fu,Vg,Oe	3.1.3.5	4OHF	2335
Fosfatasa tipo IIIB (PF03767)	Vg	3.1.3.2	3ET5	2535
Polinucleotido 5K/3-fosfatasa (PF08645)	amplia	3.1.3.32	3U7E	2001
Catión^{2+}-ATPase, tipo P (PF13246)	amplia	7.2.2.10	3B9R	16100
Fosfatasas ácidas	amplia	3.1.3.-; 3.1.3.48	2WM8	1079
Hidrolasas DHH	An,Fu,Oe,Ba,Ar	3.1.4.59; 3.1.3.7 3.6.1.1&11; 2.7.7.-	1I74, 1IR6	16058
DUF2608† (PF11019)	Oe, Ba, Ar	--	--	168
DUF705† (PF05152)	Vi, Ba	--	--	81
LNS2/lipinas (PF08235)	An,Fu,Vg,Oe,Ba	3.1.3.4	--	2000
Fosfatasa tipo factor NLI (PF03031)	An,Fu,Vg,Oe,Ba	3.1.3.16	3PGL	9785
5' nucleotidasa NT5C (PF06941)	amplia	2.7.1.113	2JAQ	1745
PGP fosfatasa (PF09419)	Fu,Vg,Oe,Ba	3.1.3.27	--	1039
fosfomanomutasa (PF03332)	An,Fu,Vg,Oe,Ba	5.4.2.8	2I55	1981
fosfatasa putativa (PF06888)	amplia	--	--	1179
Sacarosa(F)-6-P fosfatasa (PF05116)	Vg,Ba,Oe	3.1.3.24	1U2S	2067
Trehalosa-P-fosfatasa (PF02358)	amplia	3.1.3.12	1U02	7163
Pirimidín-5'-nucleotidase (UMPH-1; PF05822)	An,Fu,Vg,Oe,Ba	3.1.3.91	4NWI	1016
Semejante a hidrolasa (PF13242)	amplia	3.1.3.5	4KN8	15650
Hidrolasa 6 (PF13344)	amplia	3.1.3.3&74	2CFS	9906
Hidrolasa 3 (PF08282)	amplia	3.1.3.-	3FZQ	22502
Hidrolasas tipo-HAD (PF00702)	amplia	3.3.2.3&10; 3.8.1.3; 3.1.3.96	3WKC	58419
Tipo HAD (PF12710)	amplia	--	3FVV	15066
Tipo HAD 2 (PF13419)	amplia	EC: 3.1.3.-	3FZQ	22502

* información de PFAM, https://pfam.xfam.org/clan/HAD, consultada Feb 2020.
†DUF, dominios de función desconocida. ‡Vi, virus; Ba, Bacteria; Ar, Archea; An, Animales; Fu, Hongos; Vg, Vegetales; Oe, varios Eucariontes unicelulares.

Por otro lado, es esa misma diversidad a nivel de secuencia de aminoácidos, lo que explica la existencia de una gran variedad de funciones, algunas de las cuales, aún nos resultan desconocidas. En consecuencia, inferir la función con base en la semejanza de secuencia puede resultar engañoso.

Es sorprendente que con tal divergencia de secuencia de aminoácidos, las proteínas cl-HAD presenten una cierta conservación estructural, asociada a un mecanismo catalítico similar y que la gran mayoría sean dependientes de Mg^{2+}, aunque existen algunos ejemplos de proteínas de este grupo dependientes de Zn^{2+} y en varios casos el catión puede ser reemplazado por Mn^{2+}, Co^{2+} o Fe^{2+} y, en el caso de plantas, se han reportado enzimas que no requieren catión divalente (ver "Conocimiento de la actividad estructura y función de enzimas cl-HAD en vegetales", más adelante).

Una representación de la estructura de la haloácido dehalogenasa de *Pseudomonas* sp., la enzima que le dio el nombre a cl-HAD, se muestra en la figura 1A. Esa proteína tiene un arreglo singular y comienza en su N-terminal con una hebra β1 corta que conduce a un motivo llamado garabato (*squiggle*), semejante a una vuelta de hélice tipo π (helicoide 4_{16}). A los costados de la hebra β1 asociada por puentes de H dos motivos semejantes al plegamiento de Rosman, pero que generalmente están modificados, por lo que se ha sugerido el nombre de motivo Rosmanoide para ellos. Estos se conocen como el Rosmanoide central y el Rosmanoide C-terminal, por su posición en la secuencia de aminoácidos. El conjunto forma una lámina β-paralela que parece ser robusta y versátil. Además de que, entre el garabato y el dominio Rosmanoide central, la conexión resulta muy variable y se presentan secuencias de aminoácidos que van desde una decena para algunas, hasta cerca de un centenar como en el caso de la haloácido dehalogenasa de *Pseudomonas sp*. mostrada en la figura 1A, o más en algunos casos particulares.

En algunas fosfatasas, el garabato se conecta con el Rosmanoide central mediante un asa corta, llamada C0. En esos casos la formación de un oligómero permite que el garabato quede protegido por la interfaz y se considera que esta característica es importante en la catálisis, para reducir la exposición al solvente, pero este arreglo de asociación no covalente, otorga

© 2020 Mensaje Bioquímico. Todos los derechos reservados. ISSN-0188-137X
Comité Editorial: González Andrade, M.; Hernández Alcántara G.; Martínez González, J.J.;
Ramírez Silva, L. H. y Vilchis Landeros, M.M. *Publicado por el Departamento de Bioquímica de la Facultad de Medicina; UNAM.*

flexibilidad a la cubierta, de modo que el sustrato pueda entrar y el producto salir [8]. En la dehalogenasa de haloácido mostrada en la figura 1A, esta conexión está formada por un bulto de hélices (*helix bundle*) que hace las veces de tapa del sitio activo y que se conoce como inserción C1. Este dominio se considera una adaptación determinante de la especificidad en estas proteínas [8]. Algunos otros miembros cl-HAD tienen un dominio más extenso insertado hacia el final de Rosmanoide C-terminal, que se denomina dominio C2. La extensión C2 es a menudo un dominio de interacción proteína-proteína que permite a dichas enzimas formar oligómeros, asociarse a otros complejos multienzimáticos, o bien que puede hacer las funciones de tapa del sitio activo [8,10].

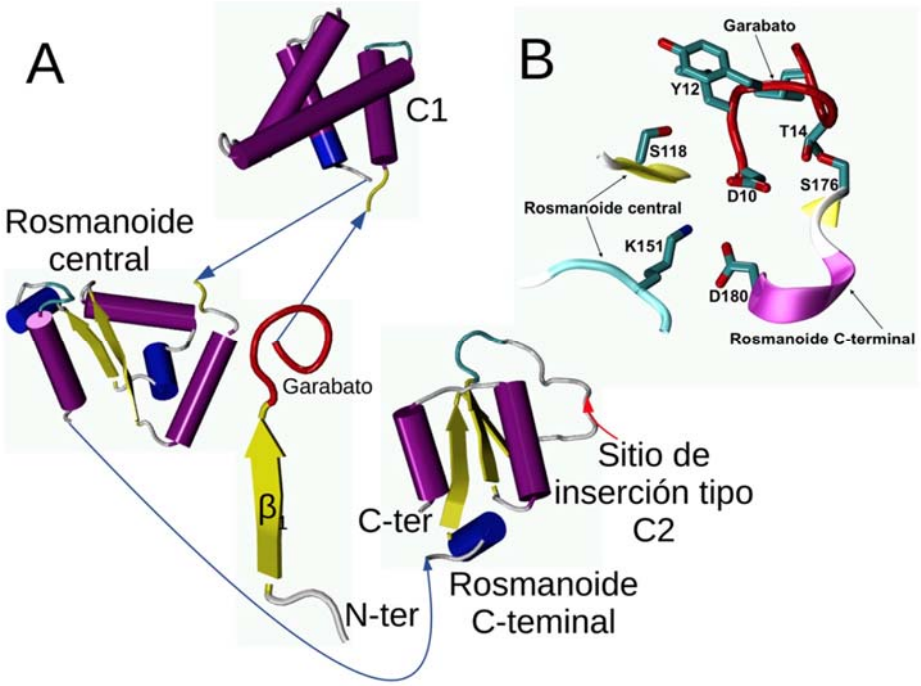

Figura 1. Estructura tridimensional de la halogenasa de haloácido (HAD) de *Pseudomonas* sp. (PDB 1jud) [7]. La representación esquemática de la estructura (**A**) se muestra dividida en los motivos clásicos del clan. Las cadenas laterales que se consideran importantes en catálisis [10] se muestran (**B**) como varillas en colores CPK. La imagen se generó usando el visualizador VMD [11].

Bases estructurales de la catálisis en las enzimas cl-HAD

La figura 1B muestra una representación de los residuos considerados importantes para la catálisis por la dehalogenasa de haloácido de *Pseudomonas* sp. Este sitio requiere la convergencia del garabato con dos segmentos del Rosmanoide central y un segmento del Rosmanoide C-terminal. En el garabato hay tres cadenas laterales que quedan orientadas para favorecer la unión del Mg^{2+} catalítico y propiciar el ataque nucleofílico del sustrato. A su vez, las cadenas provenientes de ambos constados y del lado opuesto del bolsillo contribuyen a la polarización de los grupos, lo que también contribuye a favorecer la reacción.

La figura 2, se muestra un mecanismo general de catálisis para las fosfatasas de esta familia, pero la numeración de los residuos corresponde a la dehalogenasa de haloácido mostrada en la figura 1, y además en varias fosfatasas, el residuo Y12 está substituído por un ácido aspártico.

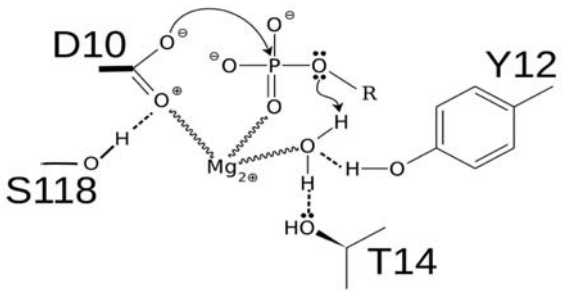

Figura 2. Representación del mecanismo general para las fosfatasas de cl-HAD. La numeración de los residuos conservados corresponde a la secuencia de la dehalogenasa de haloácido de *Pseudomonas* sp. cuya estructura se muestra en la figura 1. Variaciones en el mecanismo explican la hidrólisis de epóxidos o haloácidos, y la transferencia de fosfato [8].

© 2020 Mensaje Bioquímico. Todos los derechos reservados. ISSN-0188-137X
Comité Editorial: González Andrade, M.; Hernández Alcántara G.; Martínez González, J.J.;
Ramírez Silva, L. H. y Vilchis Landeros, M.M. *Publicado por el Departamento de Bioquímica de la Facultad de Medicina; UNAM.*

En la catálisis clásica del cl-HAD la cadena de D10 cumple el papel de grupo nucleofílico y en muchas proteínas de la familia se propone que se forma un intermediario covalente de aspartilfosfato durante la reacción (Fig. 2). Por otro lado, K151 contribuye a compensar la carga del grupo fosfato y aumenta su polarización, con ayuda del Mg^{2+} (no se muestra en la figura 1B, porque ese cristal, no tenía Mg^{2+} unido). S118 es probablemente el residuo responsable de orientar a Y12, o bien a una molécula de agua, para que se pueda donar un protón al grupo saliente. Por su parte, T14, S176, D10 y a veces el oxígeno del esqueleto de Y12 (no mostrado, por razones de claridad de la imagen) forman un sitio de unión para Mg^{2+}, cuyo papel es esencial para hacer más electrofílico al fósforo y activar a la molécula de agua que servirá para romper el intermediario aspartilfosfato, durante la segunda parte de la reacción. Es posible, que otros cationes divalentes puedan promover esta reacción, principalmente Mn^{2+}, Co^{2+} o Fe^{2+}, pero no en todas las enzimas, ya que esto depende de los demás grupos presentes en el bolsillo catalítico.

Diversidad en la especificidad de las hidrolasas de la superfamilia cl-HAD

Probablemente la caracterización más completa de las enzimas miembro de cl-HAD en un mismo organismo se realizó en la levadura del pan (*Saccharomyces cereviceae*) [12]. En este organismo ya se habían caracterizado algunas enzimas de dicha superfamilia, pero Kuznetsova y cols. [12] clonaron 13 de 19 enzimas cl-HAD solubles codificadas en el genoma de la levadura, determinaron su especificidad por diferentes sustratos y midieron los parámetros cinéticos para el mejor sustrato encontrado en cada uno de la mayoría de los casos. Además, reportan la estructura cristalina de 3 de ellas (PDB: 2QTL, 3NUQ, 3KC2 y 3RF6) y datos de otras proteínas de la levadura que interaccionaron con estas enzimas *in vitro* y que podrían representar asociaciones funcionales *in vivo*. Entre los hallazgos más significativos de este trabajo, está el papel de la enzima YKR070W, para la que encontraron una actividad de fosfatasa de baja especificidad, pero capaz de remover el fosfato de fosfopéptidos. La proteína fue ubicada en la mitocondria y su papel fisiológico no resulta aún claro, pero se piensa que sería importante en la recuperación de fosfato para la fosforilación oxidativa.

El trabajo de Kuznetsova y cols. [12] pone en evidencia la gran diversidad de sustratos fosforilados que pueden ser hidrolizados por las muchas fosfatasas cl-HAD, pero también demuestra que algunas de estas proteínas tienen baja especificidad, como YKR070W; por lo que, las mutaciones de pérdida de función en proteínas de esta superfamilia podrían ser compensadas por una actividad secundaria de otros miembros cl-HAD permitiendo al organismo sobrevivir y dar descendencia, a pesar de la mutación. En este panorama, la enzima ya mutada puede sufrir nuevas mutaciones, esta vez sin comprometer la viabilidad de la descendencia, pues la función ya estaba perdida. Eventualmente, con la acumulación de mutaciones, la proteína puede adquirir una nueva actividad que quizá resulte ventajosa y dote, al organismo que la porta, de una ventaja adaptativa. Es decir, la evolución habría generado así una nueva función favorable.

El razonamiento anterior, permite explicar, al menos en teoría, la facilidad que parecen tener las proteínas cl-HAD para evolucionar rápidamente, adquiriendo nuevas actividades y, por ende, nuevas funciones.

Funciones de estas proteínas en humanos y su relación con la salud

En una revisión reciente, se realizó una recopilación completa sobre las funciones, estructura y propiedades de diversos miembros del cl-HAD en humanos [9]. De nuevo, a pesar de la diversidad de secuencia con las proteínas ortólogas de levadura y bacterias, las enzimas del cl-HAD en humanos muestran una organización estructural semejante y en diferentes casos se encuentran presentes, alguno de los tipos de motivos de la tapa C0, C1 y C2, de modo que está representada la misma diversidad de arquitecturas encontrada en otros grupos filogenéticos. Esto indica que las diferentes clases de cl-HAD ya estaban presentes en el primer organismo metazoario primitivo, a partir del cual derivan todos los animales, incluidos los seres humanos.

Un aspecto muy interesante de las cl-HAD es su asociación con enfermedades hereditarias, que revela que algunas de sus funciones no pueden ser substituidas por otros miembros de dicha superfamilia. Es decir, a pesar de su enorme versatilidad funcional, es posible generar proteínas con funciones muy específicas, usando la misma arquitectura con la que se generan enzimas capaces de actuar sobre sustratos muy diversos, como en el caso ya mencionado en el apartado anterior para la levadura (YKR070W) [12].

Como ejemplo de su importancia podemos mencionar que las fosfatasas responsables de desfosforilar específicamente las serinas 2 y 5 del

© 2020 Mensaje Bioquímico. Todos los derechos reservados. ISSN-0188-137X
Comité Editorial: González Andrade, M.; Hernández Alcántara G.; Martínez González, J.J.;
Ramírez Silva, L. H. y Vilchis Landeros, M.M. *Publicado por el Departamento de Bioquímica de la Facultad de Medicina; UNAM.*

dominio C-terminal de la RNA polimerasa II transcripcionalmente activa, conocida como Fcp1 y que es un miembro cl-HAD. La pérdida de función de esta enzima causa el síndrome conocido como dismorfismo facial neuropático (CCFDN; *facial dysmorphism neuropathy syndrome*) [9], que representa una enfermedad autosómica recesiva, con prevalencia en familias gitanas (OMIM #604168) [13].

Otro ejemplo significativo es la enzima bifuncional polynucleotide 5'-cinasa⁄3'-fosfatasa [14] que resulta de una fusión génica de dos enzimas, siendo la fosfatasa miembro de cl-HAD. Esta enzima es esencial para que la maquinaria de reparación pueda corregir las rupturas del DNA, pues restaura el extremo 5' fosforilado y el extremo 3' desfosforilado para que la ligación pueda ocurrir [9]. Las deficiencias congénitas en esta proteína causan defectos en el desarrollo neural [15] como los que se presentan en la encefalopatía epiléptica infantil temprana (EIEE10; OMIM #613402) [16]. Por otro lado, en las células cancerosas, la sobrexpresión de esta proteína puede ayudar a los tumores a resistir la quimioterapia, por lo que resulta un blanco terapéutico de interés para el desarrollo de cocteles antitumorales [17].

Alteraciones de la función de otras cl-HAD humanas han sido implicados en el control de la hipertensión, la función cardiaca, inflamación cutánea y desarrollo del cáncer, por lo que el interés en su estudio se ha fortalecido.

Conocimiento de la actividad estructura y función de enzimas cl-HAD en vegetales

Como todos los Eucariontes, los genomas vegetales codifican para múltiples miembros de proteínas cl-HAD. Pero en el caso de las plantas, un análisis filogenético realizado por nuestro grupo de trabajo demuestra que existen genes de duplicación reciente en numerosas plantas cultivadas, tales como maíz, soya, tomate, chile, colza, quinoa y varias más. En algunos casos, sus actividades han sido estudiadas y en pocos casos la estructura ha sido resuelta.

Dada su actividad como fosfatasas, no resultó sorprendente encontrar que cuando una planta crece en suelos con poco fosfato (deficiencia mineral grave para un vegetal), la proteína cl-HAD llamada PS2 se exprese en niveles anormalmente altos. En concordancia, se reportó su actividad como fosfatasa ácida, pero que sorprendentemente, no requería Mg^{2+} [18]. Esta proteína fue clonada y reevaluada más adelante reportando que se trataba de una pirofosfatasa alcalina y, en este caso era estimulada por Mg^{2+}, pero tenía una actividad basal sin el catión [19].

Sin embargo, en una revisión reciente Gutierrez-Luna y cols. [20] hacen notar que esta enzima se parece mucho a otra proteína purificada de la planta tiempo atrás y que tenía actividad de fosfatasa ácida y de pirofosfatas alcalina dependiente de Mg^{2+}.

Nosotros hemos modelado la estructura de esta proteína usando una mezcla de métodos de modelado comparativo, simulaciones extensas de Dinámica Molecular y guiados por el índice de relevancia biológica propuesto por nuestro grupo [5]. Con el fin de proponer bases químicas para la especificidad por sustrato, estudiamos además la geometría del sitio activo, la unión a metales, la afinidad relativa por varios sustratos y la formación del intermediario fosfosapartáto, para lo cual usamos métodos semiempíricos de la Mecánica Cuántica, pero con un tratamiento completo de toda la proteína. Los resultados detallados fueron incluidos en un manuscrito en revisión [21] y aquí solo se incluyen algunos detalles interesantes y algunos datos que, por brevedad, no se trataron en el manuscrito sometido.

En la proteína PS2 de chile, encontramos un dedo de Zn^{2+} adyacente al sito activo y localizado en el motivo Rosmanoide central (Fig. 3 A). Dicho dedo de Zn^{2+} se asemeja al de una proteína bacteriana de esta familia [22], pero tanto la geometría, como la secuencia de aminoácidos de esta zona es divergente entre las dos. Esto significaría que ambas proteínas adquirieron esta secuencia de manera independiente, quizá por un evento de recombinación que se catalogaría de evolución convergente. No se ha determinado la función de este átomo de Zn^{2+} en ninguna de las dos proteínas, pero nosotros proponemos que es importante para la unión del amonio cuaternario de cholina, o el grupo con dobles enlaces deslocalizados de sustrato de la enzima bacteriana (2-hidroxi-3-ceto-5-methilthiopentenil-1-fosfato). Este último sustrato, también podría ser un sustrato alternativo en la célula vegetal. A la fecha, no se han reportado dedos de Zn^{2+} que participen en catálisis en otras enzimas, si bien son muy frecuentes en proteínas que unen ácidos nucleicos, por lo que este hallazgo resulta novedoso.

Además, la figura 3B muestra un posible intermediario predicho por los cálculos de la estructura electrónica de la proteína unida al sustrato. Para predecirlo, se simuló *in silico* la ocurrencia del ataque nucleofílico y los datos mostrados (Fig. 3B) indican que el nucleófilo no es el aspártico D11 (que equivale al D10 en la Fig 1B), sino el D13 (que

© 2020 Mensaje Bioquímico. Todos los derechos reservados. ISSN-0188-137X
Comité Editorial: González Andrade, M.; Hernández Alcántara G.; Martínez González, J.J.;
Ramírez Silva, L. H. y Vilchis Landeros, M.M. *Publicado por el Departamento de Bioquímica de la Facultad de Medicina; UNAM.*

equivale al Y12 en la Fig 1B). Además, el sitio de unión a Mg^{2+} está conformado por D11, D179 y D183 que son trés grupos carboxilato, en substitución de una serina y una treonina, dando a este grupo un exceso de carga negativa. Aparentemente, en este grupo de enzimas se ha generado una variación del sitio activo, que podría explicar cambios en especificidad inducidos por pH, como fue señalado en la revisión por Gutiérrez-Luna y cols. [20]. Por lo que, los residuos de aspartato en el sitio de unión a Mg^{2+}, son susceptibles de sufrir protonación a pH bajo, y esto llevaría a que se pierda la afinidad por el Mg^{2+}. El cambio de carga en el sitio activo resultado de la incorporación de protones podría hacer que el catión divalente ya no sea esencial, y además, ahora, la especificidad se modifique. Desde luego, hace falta una caracterización más profunda de estas proteínas, a fin de confirmar las propuestas vertidas aquí y que están soportadas por el análisis sólo realizado *in silico*.

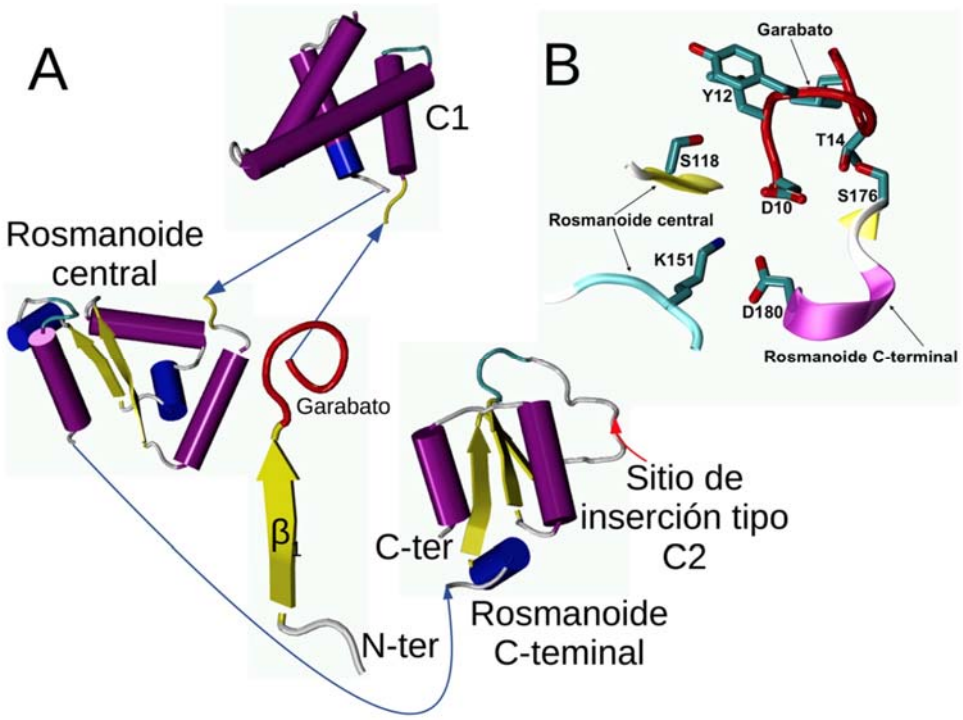

Figura 3. Predicción de los sitios de unión a metales. (A) en la proteína cl-HAD tipo PS2 de Chile (*Capsicum annuum* L.) y del intermediario aspartil-fosfato (B) con fosfocolina como sustrato.

Consideraciones finales

Un patrón común a lo largo de este capítulo es la notoria diversidad de secuencia que puede tolerar el plegamiento de los miembros cl-HAD manteniendo su plegamiento, pero cambiando la actividad y aparentemente incluso los detalles del mecanismo de catálisis. Esto las hace un ejemplo muy interesante desde el punto de vista de la ingeniería de proteínas, ya que muchas claves para entender cómo se determina la especificidad por sustrato parecen estar encerradas en la notoriamente extensa radiación evolutiva de estas proteínas.

Además, la presencia de duplicados recientes de los genes que codifican para estas proteínas en plantas, sugiere que aún continúan divergiendo y generando nuevas funciones, lo que da margen de maniobra para emplear las células vegetales como laboratorios de ingeniería de proteínas de esta superfamilia.

Por otro lado, una mejor comprensión de la relación entre su estructura y función nos podría en mejor situación para elegir miembros de esta superfamilia como blancos terapéuticos para el desarrollo de antihipertensivos, antitumorales, antiinflamatorias y otros fármacos, pues lo que hemos aprendido sobre la función de estas enzimas en los humanos sugiere que el cl-HAD tiene una gran en la fisiología humana.

© 2020 Mensaje Bioquímico. Todos los derechos reservados. ISSN-0188-137X
Comité Editorial: González Andrade, M.; Hernández Alcántara G.; Martínez González, J.J.;
Ramírez Silva, L. H. y Vilchis Landeros, M.M. *Publicado por el Departamento de Bioquímica de la Facultad de Medicina; UNAM.*

Referencias

1. Dill, K. A. y MacCallum, J. L.(2012) *Science*1042-6. DOI: http://doi.org/https://doi.org/10.1126/science.1219021
2. Martínez-Castilla, L. P. & Rodríguez-Sotres, R.On the assessment of structural protein models with ROSETTA-Design and HMMer: Value, potential and limitations Ed. H. Pérez-Sánchez; Bioinformatics Rijeka, Croatia: Intech; 2012. P. 216-238.
3. Pei, J. & Grishin, N. V. (2001) *Bioinformatics (Oxford, England)* 17, 700--12. DOI: http://doi.org/http://dx.doi.org/11524371
4. Alexander, P. A., He, Y., Chen, Y., Orban, J., & Bryan, P. N. (2009) *Proc. Natl. Acad. Sci. USA,* 21149–21154. DOI: http://doi.org/https://doi.org/10.1073/pnas.0906408106
5. Martínez-Castilla, L. P. & Rodríguez-Sotres, R. (2010) *Plos ONE,* e12483. DOI: http://doi.org/https://dx.doi.org/10.1371/journal.pone.0012483
6. El-Gebali, S., Mistry, J., Bateman, A., Eddy, S. R., Luciani, A., Potter, S. C., Qureshi, M., Richardson, L. J., Salazar, G. A., Smart, A., Sonnhammer, E. L. L., Hirsh, L., Paladin, L., Piovesan, D., Tosatto, S. C. E., & Finn, R. D. (2019) *Nucleic Acids Res.* D427-D432. DOI: http://doi.org/https://doi.org/10.1093/nar/gky995
7. Hisano, T., Hata, Y., Fujii, T., Liu, J. Q., Kurihara, T., Esaki, N., & Soda, K. (1996) J. Biol. Chem. 271, 20322-30. DOI: http://doi.org/https://doi.org/10.1074/jbc.271.34.20322
8. Burroughs, A. M., Allen, K. N., Dunaway-Mariano, D., & Aravind, L. (2006) J. Mol. Biol. 361, 1003-34. DOI: http://doi.org/https://doi.org/10.1016/j.jmb.2006.06.049
9. Seifried, A., Schultz, J., & Gohla, A. (2013) *FEBS J.* 280, 549-71. DOI: http://doi.org/https://doi.org/10.1111/j.1742-4658.2012.08633.x
10. Allen, K. N. & Dunaway-Mariano, D. (2009) *Curr. Opin. Struct. Biol.* 19, 658-65. DOI: http://doi.org/https://doi.org/10.1016/j.sbi.2009.09.008
11. Humphrey, W., Dalke, A., & Schulten, K. (1996) *J. Mol. Graph.* 14, 33-38.
12. Kuznetsova, E., Nocek, B., Brown, G., Makarova, K. S., Flick, R., Wolf, Y. I., Khusnutdinova, A., Evdokimova, E., Jin, K., Tan, K., Hanson, A. D., Hasnain, G., Zallot, R., de Crécy-Lagard, V., Babu, M., Savchenko, A., Joachimiak, A., Edwards, A. M., Koonin, E. V., & Yakunin, A. F. (2015) *J. Biol. Chem.* 290, 18678-98. DOI: http://doi.org/https://doi.org/10.1074/jbc.M115.657916
13. Congenital cataracts, facial dysmorphism, and neuropathy (CCFDN) https://www.omim.org/entry/604168, consultada en feb, 2020
14. Polynucleotide kinase 3-prime phosphatase (PNKP) https://www.omim.org/entry/605610, consultada en feb, 2020
15. Microcephaly, seizures, and developmental delay (MCSZ) https://www.omim.org/entry/6134028, consultada en feb, 2020
16. Epileptic encephalopathy, early infantile, 1 (EIEE1) https://www.omim.org/entry/308350, consultada en feb, 2020
17. Srivastava, P., Sarma, A., & Chaturvedi, C. M. (2018) *PLoS One* 13, e0190516. DOI: http://doi.org/https://doi.org/10.1371/journal.pone.0190516
18. Baldwin, J. C., Karthikeyan, A. S., & Raghothama, K. G. (2001) *Plant Physiol.* 125, 728–737. DOI: http://doi.org/https://doi.org/10.1104/pp.125.2.728
19. May, A., Berger, S., Hertel, T., & Köck, M. (2011) *Biochim. Biophys. Acta, Gen. Subj.* 1810, 178 –185. DOI: http://doi.org/http://dx.doi.org/10.1016/j.bbagen.2010.11.006
20. Gutiérrez-Luna, F. M., Hernández-Domínguez, E. E., Valencia-Turcotte, L. G., & Rodríguez-Sotres, R. (2018) *Plant Sci.* 267, 11 - 19. DOI: http://doi.org/https://doi.org/10.1016/j.plantsci.2017.10.016
21. Estrada-Gordillo, N., Valencia-Turcotte, L., Carranza-Trinidad, X., & Rogelio Rodrgíuez-Sotres (2020) *Frontiers Plant Sci.* Sometido,
22. Xu, Q., Saikatendu, K. S., Krishna, S. S., McMullan, D., Abdubek, P., Agarwalla, S., Ambing, E., Astakhova, T., Axelrod, H. L., Carlton, D., Chiu, H.-J., Clayton, T., DiDonato, M., Duan, L., Elsliger, M.-A., Feuerhelm, J., Grzechnik, S. K., Hale, J., Hampton, E., Han, G. W., Haugen, J., Jaroszewski, L., Jin, K. K., Klock, H. E., Knuth, M. W., Koesema, E., Miller, M. D., Morse, A. T., Nigoghossian, E., Okach, L., Oommachen, S., Paulsen, J., Reyes, R., Rife, C. L., Schwarzenbacher, R., van den Bedem, H., White, A., Wolf, G., Hodgson, K. O., Wooley, J., Deacon, A. M., Godzik, A., Lesley, S. A., & Wilson, I. A. (2007) *Proteins* 69, 433-9. DOI: http://doi.org/https://doi.org/10.1002/prot.21602

DR. ROGELIO RODRÍGUEZ SOTRES

Ingeniero Bioquímica Tecnológico de Monterrey y con Maestría y el Doctorado en Ciencias Químicas en la UNAM.

Realizó una estancia posdoctoral de 2 años en el King's College (Londres, Inglaterra). Es profesor titular de tiempo completo en la Facultad de Química de la UNAM desde 1993, actualmente definitivo.

Su investigación versa sobre el metabolismo vegetal, en relación con la nutrición mineral y la productividad agrícola. Pero ha colaborado con otros grupos en la investigación en estudios estructurales de proteínas y su simulación *in silico*, aplicadas a compuestos bioactivos, Ingeniería de proteínas y desarrollo de Biosensores y el modelado de mecanismos químicos de catalizadores biológicos.

Pertenece a la Sociedad Mexicana de

© 2020 Mensaje Bioquímico. Todos los derechos reservados. ISSN-0188-137X
Comité Editorial: González Andrade, M.; Hernández Alcántara G.; Martínez González, J.J.; Ramírez Silva, L. H. y Vilchis Landeros, M.M. *Publicado por el Departamento de Bioquímica de la Facultad de Medicina; UNAM.*

Bioquímica, y ha sido miembro de la Sociedad Internacional de Semillas, la Sociedad Americana de Biología de Plantas y pertenece al Sistema Nacional de Investigadores.

En la UNAM, ha sido miembro del Comité Académico del Posgrado en Ciencias Bioquímicas, Miembro del Consejo Técnico y jefe del Departamento de Bioquímica de Facultad de Química, Miembro del Comité Asesor de Supercómputo UNAM-LANCAD (DGTIC) y miembro de la Comisión Dictaminadora del Instituto de Química.

© 2020 Mensaje Bioquímico. Todos los derechos reservados. ISSN-0188-137X
Comité Editorial: González Andrade, M.; Hernández Alcántara G.; Martínez González, J.J.;
Ramírez Silva, L. H. y Vilchis Landeros, M.M. *Publicado por el Departamento de Bioquímica de la Facultad de Medicina; UNAM.*