

IDENTIFICACION DE ESTRUCTURAS TIPO LACASA (E.C. 1.10.3.2.) EN EUBACTERIA MEDIANTE Hydrophobic Cluster Análisis (HCA).

Rodríguez Buitrago J.A.
Universidad Industrial de Santander
jonbuit@yahoo.es

Recibido Enero 12 de 2006
Aceptado Mayo 15 de 2006

RESUMEN

Se implemento la metodología Hydrophobic Cluster Analysis (HCA), para evaluar la presencia y distribución de las Multi-copper Oxidasas (Lacasa EC 1.10.3.2) en eubacteria. En este estudio se expone que el análisis estructural mediante (HCA) es capaz de validar secuencias pertenecientes a la familia de proteínas en un entorno de bajo nivel de identidad, relacionando las estructuras 1D, 2D. Se presenta la secuencia Logo bacteriana y una hipótesis (Bayesiana) de relaciones evolutivas dentro de este grupo de proteínas y las secuencias capturadas en las búsquedas en bases de datos.

PALABRAS CLAVE

Lacasa, p-difenol:oxígeno-oxidoreductasa, oxidasas multicobre, Análisis de Agregados Hidrofóbicos, HCA, EC 1.10.3.2

ABSTRACT

Hydrophobic Cluster Analysis (HCA) was implemented to evaluate the presence and distribution of p-diphenol:oxygen-oxydoreductases (Lacasa EC 1.10.3.2.) in eubacteria. The structural analysis through HCA was studied showing that this is capable of validating sequences belonging to the protein family in an ambience of lower class economic level. This has been related to the structures 1D and 2D. This represents the sequence of the LOGO bacteria and a hypothesis (Bayesiana) of evolutionary relationships within this group of proteins and the sequences captured in the search for the bases of data.

KEY WORDS

Lacasa, p-diphenol:oxygen-oxydoreducataase, Multi-copper, Oxydases, Hydrophobic Cluster Analysis, HCA, EC 1.10.3.2.

INTRODUCCIÓN

Las oxidasas multicobre son proteínas involucradas en numerosas reacciones del metabolismo celular como la fotosíntesis, la fosforilación oxidativa, la homeostasis de iones metálicos y el catabolismo de nutrientes y componentes químicos tóxicos. Las lacasas (p-difenol:oxígeno-oxidoreductasas, EC 1.10.3.2) constituyen una subfamilia de enzimas que requieren O₂ para oxidar fenoles, polifenoles, metoxifenoles, aminas aromáticas y diversos sustratos no fenólicos (Decker & Terwiller, 2000). La incorporación de átomos de cobre permite a esta enzima realizar transferencia de electrones al O₂ debido a que los átomos de cobre varían su estado de oxidación entre Cu^I y Cu^{II}. Los productos orgánicos, son radicales altamente reactivos. Las lacasas participan en procesos fisiológicos como el sistema enzimático de degradación de la lignina, característico de hongos Ascomycetos y Basidiomycetos (Heinzkill et al., 1998). También están envueltas en la regulación de procesos de pigmentación de las esporas (Cárdenas et al., 2000) y patogénesis de plantas (Choi et al., 1992). Las lacasas vegetales están asociadas a procesos de polimerización de la lignina (Ranocha et al., 1999).

En la industria biotecnológica, las lacasas son usadas como biocatalizadores para diversas aplicaciones (Dean & Ericsson, 1994), tales como la degradación de lignina (Filip & Claus, 1997), el blanqueamiento de textiles (Doran & Esposito, 2000), la destoxificación de efluentes y (), la transformación de antibióticos y esteroides (Breen & Singleton, 1999).

Aunque la distribución de las lacasas se ha investigado principalmente en plantas, hongos e insectos (Mayer & Staples, 2002; Kumar, 2003), por homología de secuencia con lacasas eucariotas modelo, eg., *Coprinus cinereus*, se ha sugerido su presencia en

eubacteria (Harald Claus, 2003). No obstante, solamente en *Bacillus subtilis* y en otras pocas especies bacterianas se ha podido demostrar actividad lacasa sobre sustratos sintéticos de lacasas eucariotas (Martins et al., 2002).

Las lacasas poseen tres dominios muy similares en tamaño, todos importantes para la actividad catalítica (Enguita et al., 2003). El sitio de ligación al sustrato está localizado en una hendidura entre los dominios B y C. Un centro mononuclear de cobre está ubicado en el dominio C y un centro trinuclear de cobre está enclavado en la interfase entre los dominios A y C. El centro T1 (dominio C) sirve como aceptor de electrones del sustrato, los cuales son transferidos al centro T2/T3 (dominio A), donde el oxígeno molecular es reducido. El T1 posee un ligando axial, útil para clasificar a las lacasas de la siguiente forma: clase 1, cuando el ligando es metionina; clase 2 cuando el ligando es leucina y clase 3 cuando el ligando es fenilalanina (Neal et al., 2003).

En vista de la utilidad biotecnológica, la búsqueda de nuevas variedades de lacasas puede tener un impacto importante en las tecnologías que se sirven esta enzima. De allí la necesidad de aumentar el conocimiento sobre la distribución de las lacasas en eubacteria, terreno aún inexplorado. En esta comunicación, proveemos evidencia de una amplia distribución de las lacasas en todos los Phyla de Eubacteria, mediante el análisis de los agregados hidrofóbicos (Hydrophobic Cluster Analysis o HCA). El HCA (Callebaut et al., 1997) ha hecho posible superar las limitaciones del análisis convencional de secuencias proteicas de 1D a una en 2D, con mucho mayor poder de resolución. Por otra parte, proponemos un consenso en la estructura primaria de las lacasas bacterianas (LOGO) tal como existe para hongos y plantas (Kumar, 2003). Finalmente, mostramos una evidencia a nivel 2D de un

evento de duplicación que dio origen a todas las lacasas.

El análisis de la estructura molecular de las lacasas bacterianas servirá de plataforma para adelantar futuras investigaciones, ya que una amplia gama de nuevas enzimas permitirá la selección de mejores centros redox con aplicaciones biotecnológicas específicas, además de que facilitar el entendimiento de sus relaciones evolutivas con las eucariotas.

MATERIALES Y METODOS

Las secuencias de lacasas bacterianas fueron obtenidas de las bases de datos Genbank, Swissprot y PDB mediante búsquedas con la aplicación Psi_Blast (Altschul et al., 1997). Se buscaron homólogos de los tres dominios de lacasas modelos de *Coprinus cinereus* (SwissProt Q9Y780), *Bacillus subtilis* (Genbank AAB62305) y *Zea mays* (Genbank AAX83113) y se seleccionaron las secuencias bacterianas que presentaron los valores más bajos de identidad, en lo posible <25%. Para la determinación de las secuencias consenso se utilizó ClustalW (Thompson et al., 1994) con los parámetros por defecto (matriz BLOSUM62, gap opening=8 y gap extensión=2). Las secuencias Logo se construyeron con la aplicación WebLogo (Crook, et al., 2004) para cada uno de los centros de cobre.

El análisis estructural mediante HCA se realizó según las indicaciones de Callebaut et al., (1997), entre la familia de proteínas representativas de 7 Phyla. Las relaciones filogenéticas se construyeron utilizando los programas MrBayes v3.0B4 (Huelsenbeck & Ronquist, 2001) y PAUP (Swofford, 2002) con las secuencias que presentaron homología estructural mediante HCA en los tres dominios. Se realizaron 100.000 generaciones bajo el modelo de sustitución aminoácida de Poisson, para inferencia

Bayesiana. Como grupo externo, se usaron las secuencias de lacasas de *Manduca sexta*, *Cucurbita pepo* y *Coprinus cinereus*.

RESULTADOS

Captura de secuencias bacterianas tipo lacasa

De 190 secuencias extraídas de las bases de datos, se seleccionaron 15 como representativas de siete Phyla bacterianos con potenciales lacasas (Tabla 1). La gran mayoría, excepto *B. liqueniformes* (65%) y *O. iheyensis* (59%), están muy por debajo del umbral de identidad tolerable para relacionar estructuralmente dos secuencias (twilight zone), es decir, identidad <30% (Callebaut et al., 1997). La secuencia de más baja identidad global con las lacasas modelo fue la de *G. sulfurreducens*, perteneciente al Phylum Proteobacteria (Tabla I), la cual, no obstante la baja homología, se agrupa en el mismo clado con *B. Subtilis* (Fig. 4). El rango de identidad más amplio entre cada una de las lacasas modelo y los representantes de cada Phyla se obtuvo en *B. subtilis* (17% - 65%), seguido de *C. cinerius* (8% - 16%) y *Z. mays* (9% - 15%).

Relación estructural mediante Hydrophobic Cluster Analysis (HCA).

Los dominios de la secuencia modelo de *B. subtilis* fueron alineados a nivel 2D con 4 representantes de los Phyla Aquificaeae, Deinococcus, Cyanobacteria y Proteobacteria (Fig. 1). HCA, De esta comparación se observa que distintos motifs son conservados entre las secuencias. Fue necesario identificar los residuos aminoácidos invariantes en las estructura regulares de esta familia, desde el acercamiento de residuos "topohydrophobic", posiciones siempre ocupadas en el espacio 3D folding, por aminoácidos fuertemente hidrofobicos, (Lamarinea et al.,2001). En particular los motifs pertenecientes a estructuras secundarias regulares (hojas β y hélices α),

marcadas de la siguiente forma hojas beta de 1-25, hélices a1-3 (25 hojas β y 3 hélices α). Los cluster hidrofobicos conservados entre estas secuencias, corresponden a centros de unión al cobre presentes en los dominios 1 y 3. Los tres dominios que conforman la proteína Lacasa de *B.subtilis*, se presentan por separado, exponiendo los alineamientos de los Cluster hidrofobicos en color gris, la identidad se presenta en círculos de fondo negro y letras blancas, los sitios activos están representados por cuadros con fondo gris y letras de color blanco. Los cluster hidrofobicos determinados evidencian la homología estructural, existente entre estas secuencias huérfanas (función y/o estructura 3D indeterminada) y el aparato catalítico de la proteína cota de *B.subtilis*, se exponen en las Figura 1.

Consenso (LOGO) de los centros de cobre

Del alineamiento de 59 secuencias no redundantes, se constituyeron secuencias LOGO de los 4 centros de cobre, denominados L1 a L4 (Fig. 2). En el centro L1, se destacan como conservados los residuos, los cuales no coinciden con los consensos canónicos establecidos para centros de cobre, pero sí parcialmente con el LOGO para plantas y hongos 6/10 residuos El consenso de L2 (Tabla 2). Kumar et al., (2003) determinaron los LOGOs de las lacasas de plantas y hongos. En el alineamiento representado en secuencias logos bacteriano, se observa un patrón similar de conservación de los diferentes centros de cobre (copper Binding) característicos de todas la Lacasas, la amplia distribución de estos centros de cobre en los diferentes phyla bacterianos, la Figura 1, presenta evidencia que estas proteínas comparten aparatos catalíticos similares.

Duplicación Intraproteína: La homología intramolecular existente en la lacasas, permite afirmar que han sufrido evento de

duplicación de dominios (Kumar, 2003), el análisis (HCA) presenta evidencia de duplicación intraproteína (figura 3), en la secuencia de *B.subtilis*, lo cual indica que esta secuencia a experimentado eventos mutacionales característicos de la familia de las multi-copper oxidases. Específicamente atribuidos a lacasas eucariotas. Análisis Filogenético. Se observa en la topología que los clados poseen congruencia taxonomica a nivel de dominio. El clado conformado por las secuencias bacterianas de lacasas conduce a postular dos posibles relaciones evolutivas entre sus miembros, primero: el evento de duplicación ocurrió después del evento de especiación, con lo cual miembros del mismo phylum poseerían secuencias con mayor grado de identidad entre si que con miembros de otros phyla, por otra parte si la duplicación ocurrió antes del evento de especiación, las secuencias podrían agruparse con miembros de otros phyla como ocurre en la topología (Figura 4).

DISCUSION

En la medida en que los residuos aminoácidos intervienen en la actividad catalítica directa, éstos tienden a conservarse, de manera idéntica en todos los organismos o bien son reemplazados por residuos con las mismas propiedades eléctricas.

Alineamientos: Los sitios de unión al cobre de las secuencias bacterianas exponen igual distribución espacial que las lacasas modelos. Estos motif HXH se presentan en el dominio 1 y en dominio 3, cerca a las regiones N y C Terminal (Cullen, 1997). El alineamiento confirma que muy pocas mutaciones están ocurriendo en estas regiones de la proteína ya que son funcionalmente importantes en la familia de las multi-copper oxidases, (Harald claus, 2003). En Lacasas de origen fúngico el ligando axial del centro de unión al cobre es Leucina o Fenilalanina, en la Lacasa Bacterial de *Bacillus subtilis* el ligando axial es

Metionina al igual que para la gran mayoría de secuencias bacterianas, los residuos involucrados en la esfera de coordinación están conservados en el alineamiento. Se puede observar que las secuencias de origen bacteriano comparten valores de identidad más altos con las secuencia modelo de la lacasa fúngica (*Coprinus cinereus*) que con las secuencias modelo de lacasa de origen vegetal (*Zea mays*). La correspondencia entre los cluster HCA y elementos de la estructura secundaria ha sido demostrado estadísticamente tanto para hélices α y hojas β (Callebaut et al, 1997). La identidad entre logos bacterianos y logos de plantas y hongos (L1-L4), es del 90% y se presenta la tendencia a la conservación de motivos HXH similar a la existente en eucariotas. La existencia de residuos distintos de los presentes en los logos de plantas y de hongos se puede explicar por la variación existente dentro de este grupo de proteínas.

ESTRUCTURA 2D. El evento de duplicación expuesto en el análisis HCA para la secuencia de *B. subtilis*, permite deducir que para secuencias huérfanas dentro del rango de identidad de (11% a 59.3%) con la cota *B. subtilis*, el evento igualmente puede ser identificado. La duplicación indica que el dominio 1 dio origen al dominio 3, y que el evento de duplicación conservo la identidad de los sitios catalíticos en las secuencias huérfanas al igual que a las lacasas presentes en eucariotas. Lo anterior es consistente con la hipótesis que afirma que el evento de duplicación ocurrió basalmente en la familia de las multi-copper oxidases y que la duplicación se dio mediante domain recycling (Kumar et al., 2003). La topología obtenida mediante inferencia bayesiana, presenta un clado que relaciona las tres lacasas bacterianas con las secuencias utilizadas en el análisis, lo cual corrobora la hipótesis de pertenencia a la misma familia proteica. La ubicación de la lacasa de *Thermus thermophilus* en un subclado diferente, indica que la diversidad de las lacasas bacterianas

debe ser investigada más a fondo y que posiblemente la cota de *B. subtilis* deberá ser utilizada en conjunto con otras secuencias de posibles lacasas bacterianas como *O. iheyensis*, para mejorar el resultado de las búsquedas en bases de datos de secuencias de posibles lacasas bacterianas.

CONCLUSIONES.

Las estructuras relacionadas mediante HCA, soportan la hipótesis de origen común para las estructuras que forman parte del aparato catalítico de las lacasas eucariotas y bacterianas. La determinación de eventos de duplicación en proteínas huérfanas, similares a los eventos de duplicación encontrados en lacasas eucariotas experimentalmente caracterizadas, implica asumir que existe un folding ancestral del cual se formaron los centros de cobre de las lacasas eucariotas y bacterianas, la convergencia implicaría varios eventos evolutivos concertados que necesarios para explicar la similitud estructural evidenciada en el análisis HCA. Las perspectivas que surgen a partir de la presente investigación se enfocan en la determinación en laboratorio de la actividad lacasa de proteínas huérfanas.

AGRADECIMIENTOS

El estudiante agradece a los docentes M.Sc. Javier Jerez Jaimes por el estímulo y la colaboración en la discusión de las implicaciones evolutivas concernientes a los grupos tratados en la investigación. Al P.h.D. Jacques Chomilier por su empeño y dedicación en transmitir a estudiantes colombianos técnicas bioinformáticas que mejoran la resolución de análisis de problemas biológicos.

Tabla 1. Secuencias con posible estructura lacasa. (A).B.subtilis (B) C.cinereus (C) Z.mays

Código NCBI Accesión	Bacteria	A	B	C	Código NCBI Accesión	Bacteria	A	B	C
12722373	<i>Pasteurella multocida</i>	18	14	13	23098801	<i>Oceanobacillus iheyensis</i>	59	13	11
48780982	<i>Burkholderia fungorum</i>	19	11	10	60729683	<i>Streptomyces lavendulae</i>	22	11	11
67910359	<i>Polaromonas sp.</i>	22	13	12	48895260	<i>Trichodesmium erythraeum</i>	22	13	12
53728419	<i>Haemophilus somnus</i>	19	15	15	2983586	<i>Aquifex aeolicus</i>	20	15	11
66854867	<i>Anaeromyxobacter dehalogenans</i>	19	10	10	44004550	<i>Bacillus licheniformes</i>	65	12	11
39997751	<i>Geobacter sulfurreducens</i>	17	8	9	57238539	<i>Capmylobacter jejuni</i>	19	14	12
53795020	<i>Chloroflexus aurantiacus</i>	23	16	12	46199672	<i>Thermus thermophilus</i>	18	16	13
13473095	<i>Mesorhizobium loti</i>	19	16	13					

Tabla 2. Comparación de secuencias consensos entre eukaria y eubacteria

Centro	inicio	Consenso en eubacteria	Consenso en plantas y hongos
L1	98	TX ₂ HWHGX ₃ PX ₄ DGX	HWHGX ₉ DGX ₅ QCPI
L2	147	GTXWXH(PS)HX7QX ₃ GLXG	GTXWYHSHX ₃ QYCXDGLXGX
L3	419	HPXHLHGXXF	HPXHLHGH
L4	482	GX ₇ HCHXLHXXXGMMX	G(PA)WX(LFV)HCHIDAEXHX ₃ G(LMF)X ₃ (LFM)

REFERENCIAS BIBLIOGRÁFICAS

- Kumar S, Durani S, Prashant S, Phale E, Wangikar T. (2003). Combined Sequence and Structure Analysis of the Fungal Laccase Family, *Biotechnology and Bioengineering*, 83: 386-394.
- Daniel Cullen. (1997). Recent advances on the molecular genetics of ligninolytic fungi, *Journal of Biotechnology*, 53.pp.273-289.
- Neal T Dittmer, Haobo Jiang, Maureen J Gorman Karl J Kramer and Michael R Kanost, Richard J Suderman, Yu-Cheng Zhu. (2004). Characterization of cDNAs encoding putative laccase-like multicopper oxidases and developmental expression in the tobacco hornworm, *Manduca sexta*, and the malaria mosquito, *Anopheles gambiae*. *Insect Biochem Mol Biol*. 34: 29-41.
- Altschul, S.F., Gish, W., Miller, W., Lipman, D, Myers, E.W. (1997). Basic local alignment search tool. *J. Mol. Biol.*, pp.215, 403-410.
- Thompson D., Higgins G., Gibson J. (1994). CLUSTALW: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic acids Research*. Vol, 22, No. 22. pp. 4673-4680.
- Callebaut ., Chomilier J, Duran Labesse G, P. Poupon J., Canard I, Henrissat B. & Moron J.P. (1997). Deciphering protein sequence information through hydrophobic cluster analysis. Current status and perspectives. *Cell. Mol. Life Sci*, 53. pp. 621-645.
- M. Lamarinea, J.-P. . Berezovsky, Momon, I.N and J. Chomilier. (2001). Distribution of tightened end fragments of globular proteins statistically matches that of topohydrophobic positions: towards an efficient punctuation of protein folding?. *Cellular and Molecular Life Sciences*. 58. pp. 492-498.
- Harald Claus. (2003). Laccases: structure, reactions, distribution, *Micron* 35. pp. 93-96.
- Crooks GE, Hon G., Chandonia M. Brenner E. 2004. Chandonia JM, Brenner SE WebLogo: A sequence logo generator, *Genome Research*, 14. pp. 1188-1190.
- Huelsenbeck J. & Ronquist F. (2001). MRBAYES: Bayesian Inference of Phylogeny. *Bioinformatics*. 17: 754-755.
- Swofford D. (2002). PAUP*. Phylogenetic analysis using parsimony (* and other methods). Version 4.0b. Sinauer Associates, Sunderland, MA.
- Decker H. & Terwilliger N. (2000). COPs and robbers: putative evolution of copper oxygen binding proteins. *J. exp. Biol*. 203. pp. 1777-1782.
- Crooks GE, Hon G, Chandonia JM, Brenner SE. (2004). WebLogo: A sequence logo generator, *Genome Research*, 14: 1188-1190.
- Alexandre G & Zhulin. (2000). Laccases are widespread in bacteria. *Trends Biotechnol*, 18. pp. 41-42.
- Heinzkill M., Bech L., Halkier T., Schneider P., Anke T. (1998). Characterization of laccases from wood-rotting fungi (family Coprinaceae). *Appl. Environ. Microbiol*. 64. pp. 1601-1606.
- Cardenas W., Dankert J. 2000. Cresolase, Catecholase and laccase activities in haemocytes of red swamp crayfish. *Fish Shellfish Immunol*. 10. pp. 33-46.
- Enguita F., Martins L., Henriques A., Larrondo M. (2003). Crystal structure of bacterial endospore coat component, a laccase with enhanced thermostability properties. *J. Biol. Chem*. 278. pp. 19416-19425.
- Cardenas W., Dankert J. (2000). Cresolase, Catecholase and laccase activities in haemocytes of red swamp crayfish. *Fish Shellfish Immunol*. 10. pp. 33-46.
- Choi G. Larson T. Nuss D. (1992). Molecular analysis of the laccase gene from the chestnut blight fungus and selective suppression of its expression in an isogenic hypovirulent strain. *Mol. Plant Microbe Interact*. 5. pp. 119-128.
- Martins O., Soares M., Pereira M., Teixeira M., Costa T., Jones H., Henriques O. (2002). Molecular

and biochemical characterization of a highly stable bacterial laccase that occurs as a structural component of the *Bacillus subtilis* endospore coat, *The Journal of Biological Chemistry*, 277, pp.18849-18859.

Ranocha P, McDougall GJ, Hawkins S, Sterjiades R, Borderies G, Stewart D, Cabanes-Macheteau M, Boudet AM, Goffner D. (1999). Biochemical characterization, molecular cloning and expression of laccases-a divergent gene family in poplar. *European Journal of Biochemistry* 259, pp. 485-495.

Dean JFD Eriksson. (1994). Laccase and the deposition of lignin in vascular plants, *Holzforschung*, 48, pp.21-33.

Doran N, Esposito E. (2000). Potential applications of oxidative enzymes and phenoloxidase-like compounds in wastewater and soil treatment: a review, *Appl Cat B: Environ*, 28, pp.83-99.

Breen & Singleton. (1999). Fungal lignocellulose breakdown and biopulping, *Curr Opin, Biotechnol*, 10, pp.252-258.

Mayer A.M. & Staples R.C. (2002). Laccases: new functions for an old enzyme, *Phytochemistry*, 60, pp.551-565.