



Bioinformatics tools used in the study of phenoloxidases enzymes of the genus *Pleurotus*

Herramientas bioinformáticas usadas en el estudio de enzimas fenoloxidasas del género *Pleurotus*

Juan Zamora-Muñoz², Carmen Sánchez¹, Edna Hernández-Domínguez³, Jorge Álvarez-Cervantes⁴, Rubén Díaz^{1*}

¹Laboratorio de Biotecnología, Centro de Investigación en Ciencias Biológicas. Universidad Autónoma de Tlaxcala, Ixtacuixtla, Tlaxcala, CP. 90120, México.

²Maestría en Biotecnología y Manejo de Recursos Naturales, Universidad Autónoma de Tlaxcala, Ixtacuixtla, Tlaxcala, CP. 90120, México

³Instituto Tecnológico Superior del Oriente del Estado de Hidalgo, Apan Hidalgo, México.

⁴Universidad Politécnica de Pachuca, Zempoala Hidalgo, México.

*Corresponding author.

E-mail address: rdiazgod@hotmail.com (R. Díaz).

Article history:

Received: 21 November 2017 / Received in revised form: 29 December 2017 / Accepted: 30 December 2017 / Published online: 1 January 2018.

<https://doi.org/10.29267/mxjb.2018.3.1.95>

ABSTRACT

The bioinformatic analysis allows relating, predicting and analyzing a set of proteins of interest that share conserved areas in common by using a computer (*in silico*). This bioinformatic analysis is made from the sequence of databases, domains and key regions that make up the catalytic function of a protein. The amino acid sequence provides a prediction corresponding to the conserved regions. Subsequently, the amino acid sequences are coupled in an orderly manner by their function and structure, which generate spatial projections of tertiary structures that allow the recognition of key sequences for coordinated mutations. Bioinformatics in phenoloxidase enzymes such as laccase, manganese peroxidase and decolorizing peroxidase allows observing protein structures, alpha and beta chains as well as the ions responsible for catalysis. *In silico* techniques play an important role in different areas such as medicine, biotechnology and genomics for the benefit of human health. The DNA sequencing technique identifies organisms by means of amino

acid sequences taken from database to make template proteins, sequence alignment, homology modeling and the validation or identification of organisms.

Keywords: Amino acids, bioinformatics, phenoloxidasas, homology, proteins.

RESUMEN

El análisis bioinformático permite relacionar, predecir y analizar un conjunto de proteínas de interés que comparten áreas conservadas en común a través de un computador (*in silico*). Este análisis bioinformático se realiza a partir de la secuencia de bases de datos, dominios y de regiones clave que conforman la función catalítica de una proteína. La secuencia de aminoácidos proporciona una predicción correspondiente a las regiones conservadas. Posteriormente las secuencias de aminoácidos se acoplan de manera ordenada por su función y estructura, mismas que generan proyecciones espaciales de estructuras terciarias que permiten el reconocimiento de secuencias claves para mutaciones coordinadas. La bioinformática en las enzimas fenoloxidasa como lacasa, manganeso peroxidasa y decolorante peroxidasa permite observar estructuras proteicas, cadenas alfa y beta, así como los iones responsables de la catálisis. Las técnicas *in silico* juegan un papel importante en áreas como la medicina, la biotecnología y la genómica en beneficio de la salud humana. La técnica de secuenciación de ADN identifica organismos por medio de secuencias de aminoácidos tomadas de bancos de datos para elaborar proteínas molde, alineamiento de secuencias, modelamientos por homología y la validación o identificación de organismos.

Palabras clave: Aminoácidos, bioinformática, fenoloxidasas, homología, proteínas.

1. INTRODUCCIÓN

En el género *Pleurotus* se encontraron dos hemoperoxidasas, llamadas manganeso peroxidasa (MnP) y versátil peroxidasa (VP), además de estas peroxidadas, se detectó la producción en estos hongos ligninolíticos otra enzima fenoloxidasa denominada lacasa (Heinfling *et al.*, 1998; Mester & Field, 1998). Los hongos se clasifican atendiendo los detalles de su ciclo de vida en alguno de los cuatro filos (basidiomicetos, zigomietos, ascomicetos o quitridiomietos). La aparición de la bioinformática y la técnica de secuenciación de ADN permitieron clasificar a los hongos por medio de su ADN en donde se encuentra el material genético (Vázquez & Contreras, 2003). Las proteínas son las unidades que conforman el fenotipo y participan en la estructura y función de un ser vivo. La técnica de secuenciación permite determinar el genoma completo de una especie determinada, generando información que se almacena en bases de datos para su análisis, al conjunto de las ciencias biológicas con las herramientas computacionales se les denomina herramientas bioinformáticas (Vergara *et al.*, 2008). Dentro de las ciencias biológicas el estudio de las proteínas es muy importante gracias a estas técnicas se analiza su estructura y composición. Con la ayuda de los modelamientos es posible visualizar y generar la estructura tridimensional de las moléculas e incluso añadir interactividad y generar una aplicación que facilita el estudio en particular de algunas proteínas (Vázquez & Contreras,

2003). Algunos microorganismos ligninolíticos como los hongos de podredumbre blanca pertenecen al grupo de los basidiomicetos, este tipo de hongos se caracterizan por formar esporas sexuales llamadas basidiosporas (Sánchez, 2015).

2. PROCEDIMIENTO PARA EL ANÁLISIS BIOINFORMÁTICO DE PROTEÍNAS

En análisis bioinformático es una herramienta computacional con una serie de pasos ordenados que permite relacionar, predecir y analizar un conjunto de proteínas que comparten áreas conservadas durante su desarrollo a través de un computador (*in silico*). El análisis bioinformático puede estar constituido por una secuencia de 8 pasos, la primera etapa se conforma por la selección de la proteína de cual se tenga interés en su función biología, así como de los organismos de los cuales ha sido reportada y en donde ha sido encontrada de acuerdo a su función catalítica (Autin *et al.*, 2012). Posteriormente sigue la base estructural, en este paso se pretende describir de forma precisa la localización de los dominios a través de mapeos en la estructura de la proteína en donde la multifuncionalidad de la proteína dependerá del orden de los aminoácidos, así como de las diferentes regiones y dominios, los cuales pueden estar cercanas a la región funcional (Balasubramanian *et al.*, 2009). La tercera etapa del análisis bioinformático es la predicción a partir de secuencias de proteínas, la predicción bioinformática corresponde en la determinación de los dominios conservados que son responsables de las características físico-químicas de la proteína y donde el análisis bioinformático evalúa todas las secuencias homólogas de aminoácidos encontradas (Khan *et al.*, 2012). El cuarto paso es la creación de una base de datos con las secuencias de proteínas homólogas obtenidas en el paso anterior, esto con el fin de obtener mejores aproximaciones en la predicción de las estructuras, las bases de datos se actualizan de manera constante con la información que se obtiene de las nuevas estructuras homólogas (Hernández *et al.*, 2014). La base de datos para el análisis bioinformático se puede diseñar usando el servidor MySQL, que es un servidor bioinformático que se utiliza para generar el código Hypertext Preprocessor (PHP) y crear archivos, informes, listas y formularios que facilitan el desarrollo de las partes importantes de las estructuras homólogas, por nombrar algunos ejemplos de bases de datos se encuentra Uniprot, que es una base de datos de información de estructura, función y bibliografía sobre proteínas, o de Pfam que es una base de datos y servidor que permite identificar familias y dominios de proteínas en común de forma selectiva (Hernández *et al.*, 2014). El siguiente paso es el alineamiento, donde las secuencias homólogas son organizadas de manera óptima, en estos alineamientos se aprecian similitudes en la estructura y composición de aminoácidos. El alineamiento se da por la organización de aminoácidos de las proteínas, y nos indica los posibles cambios en las secuencias homólogas que comparten un ancestro común en la proteína de interés (Gómez *et al.*, 2008). Seguido a los pasos previos se realiza el análisis de secuencias, donde se observan motivos y dominios que sugieren un equilibrio de la secuencia, su especificidad y estabilidad termodinámica. El análisis de secuencias construye patrones y perfiles a partir de alineamientos múltiples, obteniendo regiones conservadas en un conjunto reducido de proteínas (Gómez *et al.*, 2008). El análisis de secuencias puede ser en un principio por el programa Pfam que permite identificar patrones secuenciales comunes a diversas familias de proteínas. Este programa Pfam permite buscar un conjunto restringido de dominios mediante dos subprogramas: PfamA y PfamB. PfamA proporciona una salida más

restringida, expurgada por eliminación de las dianas menos específicas. PfamB representa una salida menos restringida que contiene resultados menos específicos por lo que permite identificar posibles funciones de las proteínas que son filtradas por PfamA (Finn *et al.*, 2011). Por penúltimo paso está el modelamiento por homología, el cual genera proyecciones espaciales de estructuras proteicas a través de secuencias de aminoácidos. El modelamiento se puede realizar con la herramienta computacional Python Molecular Viewer (PMV) que proporciona el sombreado, iluminación, transparencia y reflectividad necesaria para dar mayor realismo a modelamiento (Haines, 2006). La herramienta computacional PMV modela por medio de secuencias de aminoácidos obtenidas del banco de datos Protein Data Bank (PDB) (Chen & Chen, 2002). Por último, en el análisis bioinformático de proteínas se puede realizar la generación de mutantes, que permite sugerir una coevolución o mutación dirigida de proteínas (Hernández *et al.*, 2014). El análisis de mutaciones permite la estimación de la relación evolutiva entre dos posiciones de aminoácidos en una familia de proteínas y de esta manera, identificar aquellas posiciones de los aminoácidos estructural o funcionalmente relevantes que tienden a modificarse (Hernández *et al.*, 2014).

3. BIOINFORMÁTICA DE LAS ENZIMAS FENOLOXIDASAS

Las enzimas fenoloxidasas tienen una función importante en el proceso de descomposición del carbono que es capaz de catalizar la oxidación de compuestos aromáticos, estos compuestos son degradados por las oxidasas (Varneo *et al.*, 2010). La biodegradación es un proceso complejo catalizado por enzimas que muestran inespecificidad, este proceso implica la depolimerización inicial como la transformación de los productos derivados de esta reacción para poder incorporarse en las rutas del metabolismo. El sistema enzimático de fenoloxidasas está conformado por dos tipos de actividades enzimáticas una es por oxidasas y la otra por peroxidasas, estas enzimas tienen la capacidad de actuar directamente sobre el polímero de la lignina generando en este proceso peróxido de hidrógeno (Rodríguez, 2006). La degradación de la lignina por hongos es un proceso multienzimático oxidativo en el que participan enzimas ligninolíticas y otras enzimas que son capaces de oxidar un gran número de compuestos aromáticos (Higuchi, 2004). Los hongos de pudrición blanca participan en la biodegradación de la lignina mediante mecanismos oxidativos, el sistema enzimático actúa sobre la superficie de la pared celular, las principales enzimas tienen su función catalítica directa o indirectamente sobre la lignina son la lignina peroxidasas, manganeso peroxidasas y lacasas (Díaz, 2009).

3.1. Manganeso peroxidasa

Las enzimas peroxidasas se encuentran presente en la mayoría de los hongos que degradan lignina y que puede estar presente en procesos de degradación naturales de lignocelulosa, las enzimas peroxidasa se distribuye entre los hongos de pudrición blanca (Hofrichter *et al.*, 2001). La enzima manganeso peroxidasa (MnP), se encontró por primera vez en cultivos ligninolíticos de *Phanerochaete chrysosporium*, de este organismo se han purificado diez isoenzimas nombradas H1 a H10, esto fue de acuerdo con su elución en una columna de intercambio aniónico, de ellas, cuatro correspondieron a isoenzimas de MnP (Kirk & Farrel, 1987). La enzima MnP se ha hallado en otros hongos de podredumbre blanca y en

otros basidiomicetos como *Panus tigrinus*, *Lentinula edodes*, *Nemtoloma forwardii*, *Agaricus bisporus*, *Ceriporioespora subvermispora*, *Rigidoporus lignosus* y *Trametes versicolor* (Niladevi, 2009). En las fenoloxidasas la enzima MnP (EC1.11.1.13), tiene un peso molecular aproximado de 50-60 kDa, tiene un punto isoeléctrico y un pH óptimo ácido (Hofrichter *et al.*, 2001). Esta enzima utiliza H₂O₂ para oxidar manganeso divalente (Mn II) a manganeso trivalente (Mn III) dando como resultado, pequeños oxidantes que puedan actuar sobre la lignina (Kirk & Cullen, 1998). La MnP depolimeriza compuestos fenólicos y fenoles metilados en lignina sintética, lo que indica que este sistema podría ser el que se presenta en la depolimerización de lignina *in vivo* (Kirk & Cullen, 1998). La MnP no es una enzima fuertemente oxidante y en consecuencia incapaz de atacar estructuras recalcitrantes no fenólicas predominantes en la lignina, sin embargo, si puede oxidar las estructuras fenólicas más reactivas (estas conforman aproximadamente el 10% de la lignina), estas divisiones dan origen a un grado limitado de lignolisis a través de división aril y otras vías de degradación, debido a los pequeños radicales que produce y su capacidad de penetrar la lignina, es el facilitador para el ataque posterior de otras enzimas más grandes y con mayor poder oxidante como la lignino peroxidasa (LiP) o Lacasa (Hadar *et al.*, 2012). En este trabajo la enzima MnP en estudio pertenece al organismo *Pleurotus ostreatus* con el número de acceso GenBank: AAA84397.1 formato tomado del banco de datos de National Center for Biotechnology Information (NCBI) el cual se encuentra constituido por una secuencia de 363 aminoácidos para ser modelada por homología por el programa bioinformático PMV que nos permite generar una imagen (Fig. 1), donde se puede apreciar de manera proporcional la cantidad de estructuras α -hélice y β -plegada, además en el centro superior de la enzima se aprecia el ión Mn⁺² que es utilizado por esta enzima como cofactor para su función catalítica.

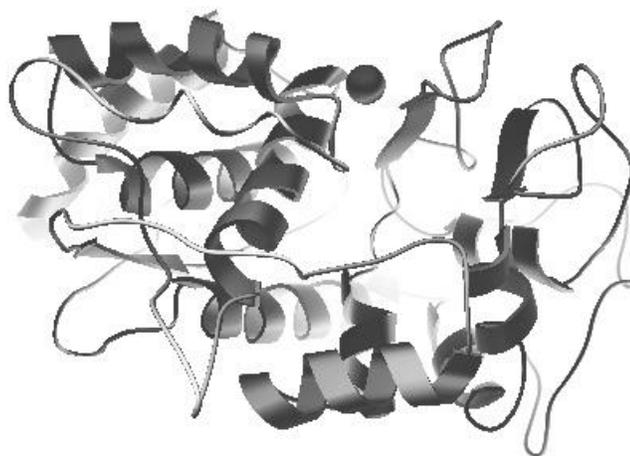


Fig. 1. Modelamiento por homologías de la enzima MnP de *Pleurotus ostreatus* (AAA84397.1) empleando el programa bioinformático PMV.

3.2. Lacasa

Las enzimas lacasa (LAC) (EC1.10.3.2), son fenoloxidasas constituidas por una secuencia de aminoácidos que difiere según el organismo al cual pertenezca. Las enzimas LAC tienen la capacidad de catalizar la oxidación de un amplio espectro de aminas aromáticas y compuestos fenólicos utilizando el oxígeno molecular como aceptor de electrones reduciéndolo a agua (Shah & Nerud, 2002). La enzima fenoloxidasa LAC contiene 4 átomos de cobre que se unen al sitio activo y son los responsables de la oxidación del sustrato, esta enzima se encuentra ampliamente distribuida en hongos, plantas, bacterias e insectos y en cada uno de los casos su función es diferente (Gianfreda *et al.*, 1999). Son glicoproteínas extracelulares con pesos moleculares entre 20 y 80 kDa y aproximadamente del 15 al 20% de su peso molecular está dado por carbohidratos (Díaz *et al.*, 2013). Se han clasificado dentro del grupo de las enzimas denominadas azul cobre oxidasas, estas enzimas requieren oxígeno y cobre para llevar a cabo su función oxidativa (Claus, 2003). Las enzimas LAC a diferencia de la MnP, no requiere peróxido de hidrógeno (H_2O_2) para oxidar al sustratos (Cordoba & Cultid, 2015). El sitio activo de estas enzimas está formado por cuatro átomos de cobre, un átomo del tipo T1, un átomo del tipo T2 y dos átomos del tipo T3, los sitios de los dos iones de cobre T3 están antiferromagnéticamente acoplados, los sitios T3 y T2 están organizados en un único grupo que es capaz de ligarse al oxígeno, el cual es el aceptor final de electrones (Niladevi, 2009). Las LAC tiene una coloración azul dada por un átomo de cobre, las enzimas que no presenta este átomo de cobre se les denominan LAC blancas o amarillas (Martínez, 2002). El ion cobre tipo T1 es responsable del color azul de la proteína, la cobre tipo T2 se caracteriza por no presentar absorbancia detectable en la región del visible y por tener una alta afinidad por aniones que actúan como inhibidores de la actividad de la enzima, la cobre tipo T3 son un complejo binario formado por un par de iones unidos por puentes de hidrógeno con un máximo de absorbancia de 330 nm y un espectro de fluorescencia típico (Duran, 2007). Para este trabajo la enzima LAC en estudio está constituida por una secuencia de 533 aminoácidos y pertenece al organismo *Pleurotus ostreatus* con el número de acceso GenBank: AAR21094.1 (NCBI) para ser modelada por homología por el programa bioinformático PMV, donde se puede apreciar una mayor cantidad de estructuras β -plegada y una menor cantidad de estructuras α -hélice que contiene en su modelamiento, asimismo, se puede observar el núcleo trinuclear de átomos de cobre en el sitio activo de la enzima (Fig 2).



Fig. 2. Modelamiento por homología de la enzima LAC de *Pleurotus ostreatus* (AAR21094.1) empleando el programa bioinformático PMV.

3.3. Decolorante peroxidasa

La enzima decolorante peroxidasa (DyP) (E.C. 1.11.1.19) tiene la capacidad de degradar un amplio espectro de colorantes, se encuentran distribuidas de forma extendida al igual que otras peroxidadas, estas proteínas en su mayoría se encuentran presentes en bacterias, mientras que solo una pequeña cantidad se encuentra en hongos y eucariotas superiores. La familia de estas enzimas están conformadas por unos cuatro mil miembros de los que 3707 son bacterias, a los eucariotas les corresponden 117 miembros y solo se encuentran 11 miembros en archa, esto puede sugerir que se trata de la contraparte bacteriana de las peroxidadas presentes en hongos (Fraaije & Bloois, 2013). Las enzimas decolorante peroxidasa son bifuncionales ya que aparte de la actividad oxidativa que presentan también aportan actividad hidrolítica (Colpa *et al.*, 2014). Estas enzimas forman parte de la familia de las hemoperoxidasas de clase II que han evolucionado de manera independiente (Hofrichter *et al.*, 2010). Son hemoproteínas extracelulares que han sido descubiertas recientemente, estas enzimas tienen la capacidad de oxidar varios colorantes particularmente los xenobióticos que derivan de las antraquinonas, al igual que compuestos fenólicos (Sugano, 2009a; Sugano, 2009b). Una característica importante es su capacidad para oxidar los colorantes de las antraquinonas ya que estos colorantes pueden ser oxidados por otras pocas peroxidadas y es una característica importante de las DyP (Kim & Shoda, 1999). Estas enzimas forman parte de un sistema de biocatálisis alternativo para la oxidación directa de compuestos aromáticos recalcitrantes metoxilados (Liers *et al.*, 2010). El pH óptimo de estas enzimas se encuentra entre 2 y 4, además tiene una amplia estabilidad térmica y una tolerancia relativamente alta a concentraciones de H₂O₂ (Salvachúa *et al.*, 2013). En la figura (Fig.3) se muestra a la enzima DyP de *Pleurotus ostreatus* con el número de acceso GenBank: CAK55151.1 (NCBI), la cual contiene 525

aminoácidos. Se puede apreciar una cantidad similar de estructuras β -plegadas y estructuras α -hélice.



Fig. 3. Modelamiento por homologías de la enzima DyP de *Pleurotus ostreatus* (CAK55151.1) empleando el programa bioinformático PMV.

4. LA BIOINFORMÁTICA UNA HERRAMIENTA MULTIDISCIPLINARIA Y SUS APLICACIONES EN LA SOCIEDAD

La bioinformática es una disciplina que ocupa tecnologías para analizar, organizar y distribuir información biológica y poder responder preguntas sobre las ciencias biológicas. La bioinformática es una herramienta multidisciplinaria la cual puede ser descrita como la interacción conjunta entre la computación y la biología, en donde esta disciplina ayuda a mejorar la calidad de vida humana, las condiciones físicas, además de seguir disipando incógnitas del genoma humano dando paso a una nueva era de la investigación genómica (Zepeda, 2013) . Esta disciplina ha tenido grandes avances en la actualidad como la detección y tratamiento de enfermedades y la producción de alimentos genéticamente modificados por mencionar algunos ejemplos. La bioinformática involucra la solución de problemas complejos que incluyen la colección, organización, almacenamiento y recuperación de la información biológica que se encuentran en las bases de datos (Zepeda, 2013). De acuerdo con el NCBI, la bioinformática es un campo de la ciencia en donde la biología, la computación y las tecnologías de la información se integran con el fin de facilitar el descubrimiento de nuevas ideas biológicas y fortalecer los principios unificadores en la biología. La bioinformática crea y mantiene bases de datos donde se almacena información biológica como secuencias de nucleótidos y aminoácidos. En la actualidad la bioinformática es el campo más popular en el análisis e interpretación de estructuras proteicas, secuencias de nucleótidos, aminoácidos y dominios de proteínas, al proceso de analizar e interpretar los datos se conoce como biocomputación. La biotecnología y la medicina molecular con ayuda de la biocomputación y la bioinformática

pretenden incrementar la investigación genómica y post genómica haciendo uso de estas herramientas bioinformáticas (Zepeda, 2013).

5. LA BIOINFORMÁTICA COMO HERRAMIENTA EN LA CLASIFICACIÓN DE ORGANISMOS

En la última década se han desarrollado métodos experimentales que permiten estudiar diferentes fenómenos en las ciencias biológicas, de manera que es posible conocer los genes que componen el genoma de cualquier organismo, esto es posible gracias a una técnica llamada secuenciación de ADN (Sanger *et al.*, 1977). La aplicación de la secuenciación y otras técnicas han proporcionado en los últimos años grandes cantidades de información que se almacenan en bases de datos, el análisis de estos datos así como el modelamiento por homología han aparecido como pasos secuenciales en las ciencias biológicas (Vergara *et al.*, 2008). La clasificación de los hongos se realiza por su morfología y su ciclo de reproducción, las herramientas bioinformáticas y la técnica de secuenciación de ADN, clasifican a los hongos en alguno de los filos como basidiomicetes, zigomietes, ascomicetes, quitridiomietes o deuteromicetes (Bon, 2005). El análisis y la aplicación de las herramientas bioinformáticas ha podido clasificar al género *Pleurotus* en el filo de los basidiomicetos (Bon, 2005). Como se puede observar en la figura (Fig. 4a), el primer paso para la secuenciación de ADN es la síntesis de una molécula de ARN usando como molde un segmento de una de las cadenas de ADN, al transcribirse un gen en el núcleo de la célula, el ARN contiene las regiones de los intrones como las de los exones, esta molécula sufre un proceso de maduración llamado “*splicing*” donde se eliminan los intrones que son no codificantes, generando de esta manera un RNAm maduro (Fig. 4b), para luego ser exportado a los ribosomas donde se traduce la información y se obtienen las proteínas. Una vez obtenida la secuencia de aminoácidos correspondientes a una proteína específica, se puede realizar trabajo bioinformático para obtener modelamientos de las estructuras, como se observa en la figura (Fig. 4c) donde la secuencia específica se introduce a un ordenador para dar origen a la posible estructura tridimensional de una proteína (Necochea & Canul, 2004). Las proteínas se pueden modelar a partir de la estructura primaria como se muestra en la figura (Fig. 4d), las proteínas se conforman en estructuras tridimensionales complejas, esto se debe a que la información contenida en la estructura primaria dicta un número de diferentes conformaciones posibles y que aquella más estable termodinámicamente es la conformación nativa y funcionalmente activa, así de esta manera se obtiene la predicción y modelamiento computacional de la estructura de las proteínas (Fig. 4e). Posteriormente, pueden utilizarse las herramientas bioinformáticas para la identificación de organismos a partir de la secuenciación de ADN, en la figura (Fig. 4f) se muestra como ejemplo el cuerpo fructífero de *Pleurotus* como organismo identificado por la técnica de secuenciación de ADN (Anfinsen, 2008).

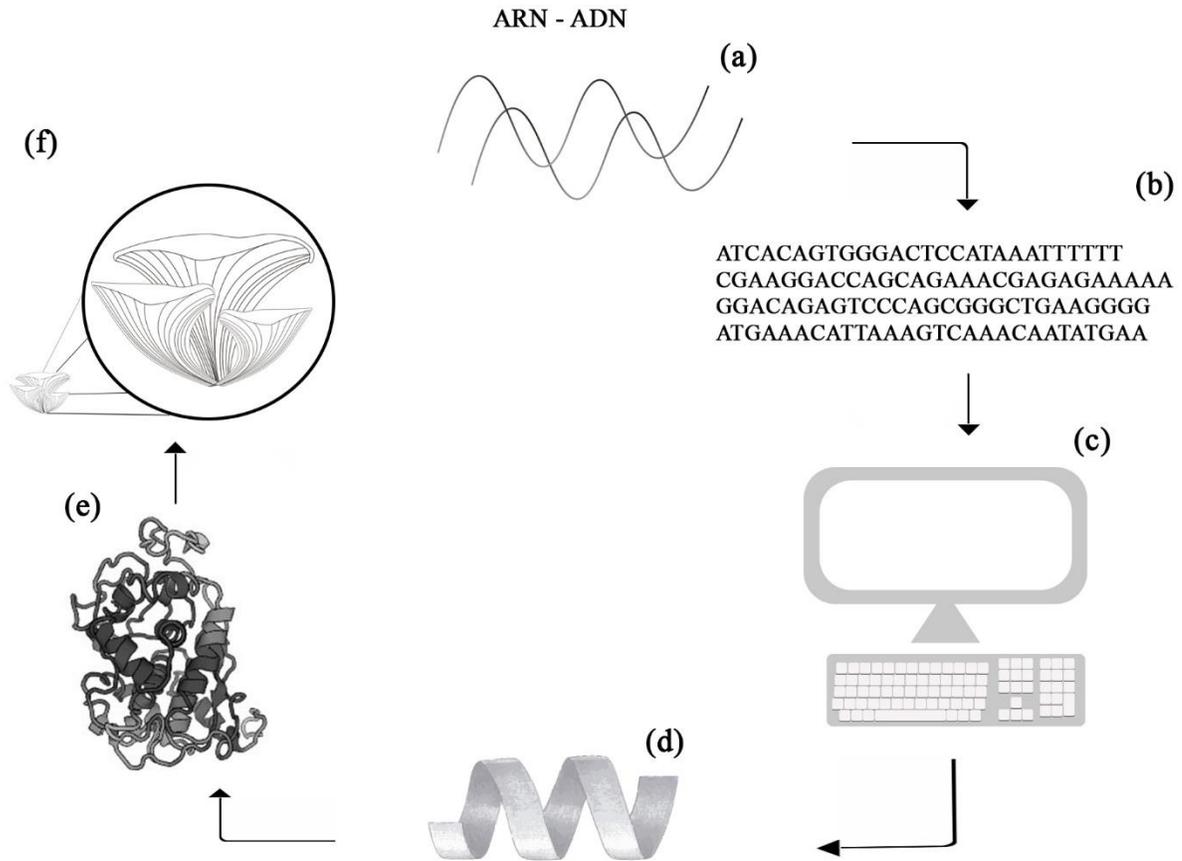


Fig. 4. Secuenciación de ADN, (a) transcripción de genes, (b) ordenamiento lineal de las secuencias, (c) análisis a través de un ordenador, (d) modelamiento de la estructura primaria, (e) modelamiento de estructuras complejas e (f) identificación de organismos.

5.1. Diseño bioinformático en 3D por secuencias de aminoácidos

Los diseños estructurales en 3D de la mayoría de las proteínas no se han determinado y muchas de ellas no tienen un ligando conocido, para estas situaciones ni los métodos basados en la estructura, ni los métodos basados en ligandos se puede ocupar para describir el diseño, por lo que se necesita un método para predecir las interacciones proteína ligando, en ausencia de estructuras o información del ligando. Recientemente se diseñó un modelo basado en secuencias para las interacciones proteína ligando construido sobre la secuencia primaria de las proteínas y las características estructurales de moléculas pequeñas utilizando el enfoque de la máquina de vectores de soporte (Wang, 2008). El modelo de vectores es una técnica que fue aplicado usando 15000 interacciones proteína ligando entre 626 proteínas y más de 10 mil compuestos activos (Lemer *et al.*, 2008). Para el diseño de modelos de secuencias conocidas se utilizan las diferentes bases de datos de secuencias genéticas que cuentan con una amplia disponibilidad de secuencias de ADN y secuencias de aminoácidos, en la tabla 1, se pueden observar las bases de datos que intercambian

información sobre secuencias producidas en los laboratorios, procedentes de más de 100000 organismos distintos.

Tabla 1. Bancos de datos bioinformáticos utilizados en modelamiento por homología de proteínas.

| Nombre | Descripción | URL | Referencia |
|-------------------|--|---------------------------------|--------------------------------|
| SRS | Sequence Retrieval System | srs.ebi.ac.uk | Kreil & Etzold, 2003 |
| CATH | Class-Architecture-Topology/fold-Homologus superfamily | www.biochem.ucl.ac.uk/bsm/cath/ | Orengo <i>et al.</i> , 1999 |
| MODBASE | | guitar.rockefeller.edu/modbase/ | Sánchez & Sali, 2003 |
| PDB | Protein Data Bank | www.rcsb.org/pdb/ | Berman <i>et al.</i> , 2000 |
| SWISS-PROT | Swiss Protein | www.ebi.ac.uk/swissprot | Gasteiger <i>et al.</i> , 2001 |
| GenBank | Base de datos de secuencias de genes | www.ncbi.nlm.nih.gov/GenBank | Benson <i>et al.</i> , 2000 |

5.2. Diseño bioinformático en 3D por homología

El diseño bioinformático para el modelamiento de proteínas es realizado a través de secuencias de aminoácidos, que son organizadas y validadas con ayuda de las herramientas tecnológicas para generar una imagen en 3D de una proteína modelada por homología. La proteína obtenida es catalogada como una innovación en el concepto de evolución molecular, este modelo parte de la semejanza que hay entre las estructuras primarias de las proteínas y de proteínas homólogas de estructuras tridimensionales conocidas (proteínas molde), esto implica directamente en la similitud estructural que hay entre ellas (Chothia & Lesk, 2003). Como se muestra a continuación (Fig. 5) los métodos de modelado de proteínas por homología se estructuran básicamente por una serie de pasos que son la identificación y selección de proteínas modelar (Fig. 5a), posteriormente sigue la alineación de las secuencias obtenidas del organismo en estudio, así como su número de identificación (Fig. 5b), para seguir con el alineamiento de las regiones conservadas de las secuencias de aminoácidos de cada proteína (Fig. 5c) y de las cadenas moldes como las de los residuos de aminoácidos responsables del sitio activo (Fig. 5d), estos pasos son seguidos de la construcción de las coordenadas del modelo y la validación (Fig. 5e), por último se proyecta una estructura modelada por homología (Fig. 5f). Este proceso es llevado a cabo

por los programas y servidores de la red mundial que son las herramientas necesarias que proporcionan los datos existentes en las diferentes bases de datos (Santos & Bicca, 2003). El modelado y diseño por homología indica las relaciones funcionales y evolutivas en las proteínas ya que pueden poseer secuencias de aminoácidos similares, así como numerosos restos idénticos o químicamente semejantes en posiciones idénticas y que estas estructuras están estabilizadas por las diferentes interacciones electrostáticas, hidrofóbicas, puentes de hidrógeno o enlaces disulfuro (Trudy & Mckee, 2003).

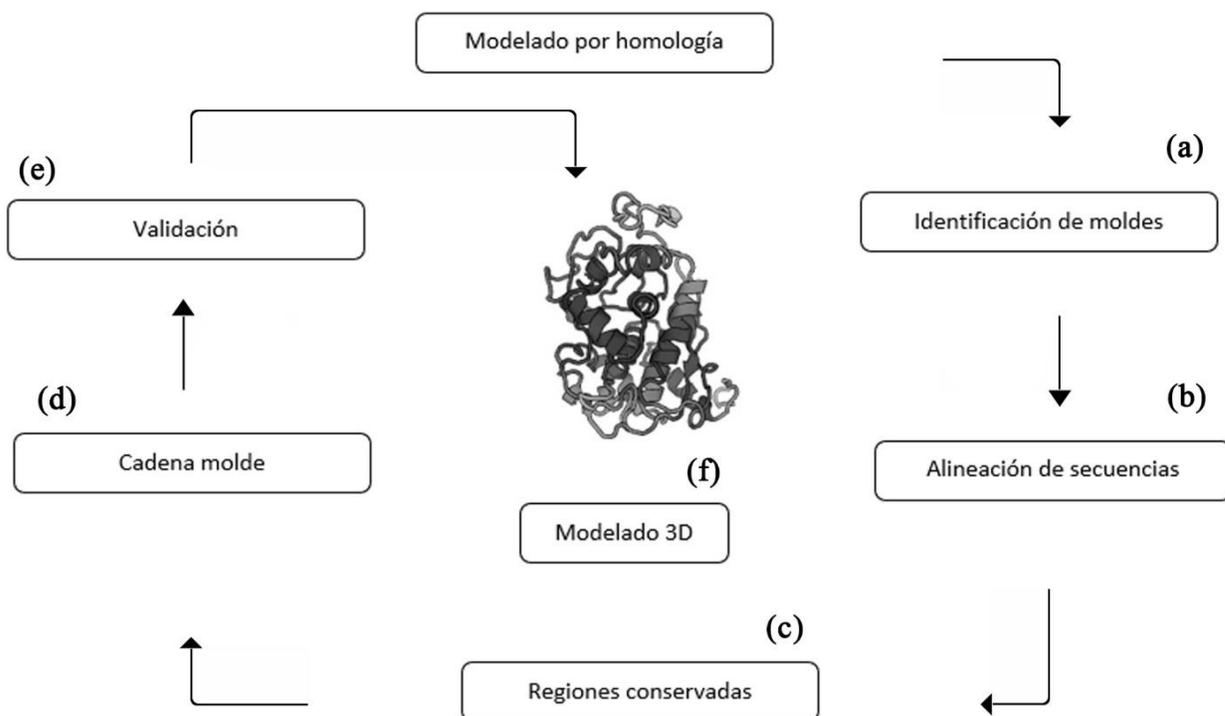


Fig. 5. Modelamiento de proteínas por homología, (a) identificación de moldes, (b) alineación de secuencias, (c) regiones conservadas, (d) coordenadas del modelo, (e) validación de la proteína modelada (f) obtención del modelado por homología.

5.3. Proteínas molde en el diseño del modelado por homología

Una proteína molde es una proteína que tiene una estructura tridimensional con una posición afín a la secuencia de aminoácidos que se utilizará de molde para la proteína de estudio (Peitsch, 1993). A través de la proteína molde que se emplea es posible conocer o no a la familia proteica a la que pertenece la proteína de estudio; para el primer caso se selecciona la proteína molde directamente del servidor PDB, si se desconoce la proteína de estudio se busca de forma sistemática uno o más moldes adecuados en una base de datos de estructuras primarias derivadas de las proteínas almacenadas en el mismo servidor, además se pueden utilizar las herramientas BLAST36 o el FastA37 de las bases de datos de secuencias de pares de bases de ADN o de aminoácidos (Barker *et al.*, 2000). Si el grado de identidad entre las estructuras primarias de las proteínas molde y de la proteína problema es igual o superior del 25% y el número de residuos es superior a 80, puede existir una gran probabilidad de que estas proteínas tengan estructuras tridimensionales y se puede construir un modelo para la proteína problema, si se llegara a tener una similitud entre las secuencias menor o cerca del 25% se tienen que utilizar métodos de compatibilidad de secuencia y estructura o modelos (Jones & Thornton, 1993). El método BLAST36 Y FastA37 evalúan el grado de igualdad entre las estructuras terciarias de la proteína en estudio y de las posibles proteínas molde cuya estructura terciaria es más adecuada en relación a la secuencia de la proteína estudiada, la evaluación de estructura de la secuencia se realiza a través de un potencial empírico derivado de una tabla de residuos observados en las proteínas de estructuras conocidas (Rost & Sander, 1996). En la tabla 2 se pueden observar los diferentes servidores bioinformáticos que sirven como herramientas computacionales para la validación de modelos y el modelado de las proteínas por homología.

Tabla 2. Servidores bioinformáticos utilizados para el modelamiento por homología de proteínas.

| Nombre | Descripción | Función | URL | Referencia |
|--------------------|--|------------------|--|------------------------------|
| BIOTECH | | Valida modelos | biotech.embl-heidelberg.de:8400/ | Hooft <i>et al.</i> , 1996 |
| PROBE | Estabiliza los datos de codificación robusta | Valida modelos | www.ucmb.ulb.ac.be/UCMB/PROVE | Pontius <i>et al.</i> , 2003 |
| SWISS-MOD | Modela proteínas por homología | Modela proteínas | www.expasy.ch/swissmod | Peitsch & Jongeneel, 1993 |
| SWISS-MODEL | Modela proteínas por comparación | Modela proteínas | www.expasy.org/swissmod/SWISS-MODEL.html | Guex & Peitsch, 1997 |
| BLAST | Alinea | Alinea | www.ncbi.nlm.nih.gov/BLAST | Altschul <i>et</i> |

| | | | | |
|-----------------------|--|--|--|---------------------------------|
| | secuencias de ADN, ARN o de proteínas | secuencias | / | <i>al.</i> , 1997 |
| CLUSTAL | Alinea múltiples secuencias | Alinea secuencias | www.ebi.ac.uk/clustalw/ | Jeanmougin <i>et al.</i> , 1998 |
| FastA | Representa secuencias usando códigos de una única letra | Alinea secuencias | www2.ebi.ac.uk/fasta3 | Pearson, 1990 |
| MultAlin | Múltiple sequence alignment | Alinea secuencias | prodes.toulouse.inra.fr/multalin/multalin.html | Corpet, 1988 |
| BLAST | Alinea secuencias | Provee estructuras moldes | www.ncbi.nlm.nih.gov/BLAST/ | Altschul <i>et al.</i> , 1990 |
| FastA | Representa secuencias usando códigos de una única letra | Provee estructuras moldes | www2.ebi.ac.uk/fasta3 | Pearson, 1990 |
| UCLA-DOE-FRSVR | Provee información de las ciencias de la salud y el medio ambiente | Provee de estructuras moldes | fold.doe-mpi.ucla.edu/ | Fischer & Eisenberg, 1996 |
| ExpASySIB | Expert Protein Analysis system | Provee datos de bioinformática general | www.expasy.org/ | |

5.4. Alineación de secuencias en el modelado de proteínas

En la alineación de la secuencia en estudio con la secuencia molde es necesario alinear los residuos estructuralmente equivalentes tomando en cuenta las características estructurales que se tienen en común, tales como elementos de estructura secundaria y residuos catalíticos (Santos, 2000). Durante el proceso de alineamiento se generan espacios vacíos denominados “gaps” que son representados en la alineación por líneas discontinuas que principalmente corresponden a las secciones de la secuencia que se han perdido a través del

tiempo, durante la identificación de moldes se tiene que alinear la secuencia de estudio con las secuencias de todos los moldes disponibles, en este caso se obtiene una alineación múltiple. La alineación múltiple se considera más confiable que la alineación simple que implica solo dos secuencias, la razón es que la alineación múltiple permite detectar más fácilmente las características estructurales comunes de proteínas homólogas en donde la mayoría de los métodos de alineación se basan en técnicas de programación dinámicas (Smith & Waterman, 2003). Después del alineamiento es posible reconocer regiones conservadas y regiones variables, en este caso las regiones conservadas corresponden a las zonas de máxima similitud en donde las conformaciones son muy similares, por otra parte, en las regiones variables no hay correspondencia estructural, la mejor alineación de varias secuencias conocidas se obtienen por superposición de las moléculas o de las restricciones espaciales. Con la superposición se mejora la calidad de la alineación con ayuda de otras informaciones como cuando se tienen varias secuencias homólogas se puede construir un perfil estructural de la familia proteica (Rychlewski *et al.*, 1998). Además, se puede utilizar información proveniente de las estructuras que fueron modeladas como guía para la alineación, la información permite modificar de manera interactiva la alineación y por lo tanto la estructura tridimensional de la proteína en estudio (Sánchez & Sali, 2003). En la tabla 3 se pueden observar los diferentes programas bioinformáticos que sirven como herramienta computacional para el modelado de estructuras molde, la validación de modelos y el modelado de las proteínas por homología que se pueden utilizar.

Tabla 3. Programas bioinformáticos utilizados para el modelado por homología de proteínas.

| Nombre | Descripción | Servicio | Dirección | Referencia |
|-----------------------|---|-------------------------------------|--|--------------------------------|
| PROCHECK | Comprueba la calidad estereoquímica | Valida los modelos | www.biochem.ucl.ac.uk/roman/procheck/procheck.html | Laskowski <i>et al.</i> , 1993 |
| WHATCHECK | Evalúa el plegamiento de proteínas | Valida los modelos | www.cmbi.kun.nl/swift/whatcheck/ | Hooft <i>et al.</i> , 1996 |
| MODELLER | Modela y compara estructuras tridimensionales | Modela proteínas | guitar.rockefeller.edu/modeler/.html | Šali & Blundell, 1993 |
| SwissPdbViewer | Analiza varias proteínas al mismo tiempo | Modela proteínas | ca.expasy.org/spdbv/ | Guex & Peitsch, 1997 |
| PROFIT | Predice la estructura de proteínas | Genera fuentes de estructuras molde | www.bioinf.org.uk/software/ | Lathrop <i>et al.</i> , 1995 |

5.5. Elaboración de modelos tridimensionales empleando herramientas bioinformáticas

El modelado de las regiones conservadas estructuralmente se realiza en un espacio tridimensional en donde el modelamiento se puede apreciar en sus tres dimensiones (3D), de esta manera se localiza el mayor número de átomos posibles de la proteína en estudio de la cual depende la alineación (Peitsch, 1997). La construcción de la proteína en estudio está basada en la conformación de la cadena principal de la cadena molde la cual puede ser transferida a ella. En la actualidad hay muchos métodos de modelado de las regiones conservadas, unos de los más importantes son el modelado por la unión de cuerpos rígidos, el modelado por la combinación de segmentos y el modelado por la satisfacción de restricciones especiales (Aszódi & Taylor, 1997; Levitt, 2003; Greer, 1990). El modelado de las regiones conservadas tiene diferencias funcionales entre miembros de una misma familia proteica, esto es en consecuencia de las diferencias estructurales en la superficie externa de las proteínas, esto es por las sustituciones, eliminaciones e inserciones de los residuos en las cadenas de proteínas homólogas, al ser regiones estructuralmente variables las zonas conservadas determinan la especificidad de las proteínas, contribuyendo a la estructura de los sitios de conexión y para el caso de enzimas es el sitio activo por lo que el modelado es una etapa esencial en la generación de modelos proteicos por homología (Fiser *et al.*, 2000).

5.6. Validación de proteínas modeladas por homología

La validación de un modelo obtenido puede ser realizado en diferentes niveles de organización estructural, se tiene que evaluar la calidad del empaquetado global de la proteína, los probables errores estructurales en regiones localizadas y los parámetros estereoquímicos (Johnson *et al.*, 2003). La calidad del modelo obtenido depende de la estructura de la proteína molde utilizada, se necesita que esta última haya sido obtenida con buena resolución, una característica de que una proteína fue modelada correctamente es que debe tener una estructura terciaria termodinámicamente estable. La calidad del modelado también depende de la proteína elegida como molde y de la alineación realizada, es muy importante comprobar si se encuentran diferencias conformacionales entre los elementos de la estructura secundaria de las regiones conservadas de las estructuras molde y de la estructura modelada (Santos *et al.*, 2001). Uno de los programas más utilizados para la evaluación de los parámetros es Procheck, el cual evalúa las longitudes de ligación, la planicidad de los anillos de cadenas laterales, los ángulos planos, la quiralidad, la planicidad de las conexiones peptídicas, la conformación de las cadenas laterales, los ángulos de la cadena principal y de las cadenas laterales (Ver programa bioinformático en Tabla 3), el impedimento estérico entre pares de átomos no deseados y la calidad del modelo generado (Ramachandran & Sasisekharan, 2003; Laskowski *et al.*, 1993). El programa Whatcheck brinda información acerca de la formación de regiones centrales hidrofóbicas, la accesibilidad de residuos y la distribución espacial de grupos iónicos, la distribución de las distancias atómicas y las conexiones con el hidrógeno de la cadena principal (Ver programa bioinformático en Tabla 3) (Hooft *et al.*, 1996). Por otra parte, el programa Prosa II tiene la capacidad de evaluar el ambiente de cada aminoácido de la proteína modelada, tomando como referencia el ambiente esperado de las proteínas

análogas de alta resolución (Sippl, 1993). Existen programas en donde se encuentran otras técnicas basadas en mecánica molecular y cálculos de energía libre de solvatación, estas técnicas son capaces de estimar la calidad de la estructura terciaria de los modelos de proteínas y la validación del modelamiento por homología (Sauder *et al.*, 2000). La validación de los modelamientos de proteínas por el programa bioinformático PMV de las enzimas MnP, LAC y DyP en este escrito pueden ser validados a través de la relación entre la estructura y función de un modelo construido a partir de secuencias de aminoácidos, para esto se tomara como correcto si el largo de la cadena es <150 aminoácidos, si el Root-Mean-Square-Deviation (RSMD) global sobre el carbono alfa es <7.0 Angstroms, si los porcentajes de los carbonos alfa son equivalentes a >30% y si el porcentaje de identidad de secuencia en el alineamiento utilizado para construir los modelos es <40%.

6. CONCLUSIONES

Las herramientas bioinformáticas permiten realizar modelamientos por homología y la identificación de organismos por medio de la secuenciación de ADN. A través de las técnicas bioinformáticas se obtienen árboles filogenéticos, regiones conservadas, estructuras molde, bases de datos, dominios de los sitios activos y mutaciones coordinadas del organismo de estudio (*in silico*). El uso de las herramientas bioinformáticas es una alternativa potencial para el estudio de los fenotipos y genotipos de las proteínas y organismos de interés.

CONFLICTO DE INTERESES

Los autores declaran que no existe conflicto de intereses.

REFERENCIAS

- Altschul S., Madden T., Schaffer A., Zhang J. & Miller W. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research*. 25(17): 3389-3402.
- Aszódi A. & Taylor W. 1997. Distance geometry based comparative modelling. *Folding and Design*. 1(5): 325-334.
- Altschul S., Gish W., Miller W., Myers E. & Lipman D. 1990. Basic local alignment search tool. *Journal of Molecular Biology*. 3(10): 215- 403.
- Anfinsen C. 2008. Principles that govern the folding of protein chains. *American Association for the Advancement of Science*. 181(96): 223-30.
- Autin L., Johnson G., Hake J., Olson A. & Sanner M. 2012. Upy: A Ubiquitous computer graphics Python API with Biological Modelling Applications. *Computer Graphics and Applications, Institute Electrical and Electronics Engineers*. 32(5): 50-61.
- Balasubramanian S., Kannan T., Hart P. & Baseman J. 2009. Amino acid changes in elongation factor of *Mycoplasma pneumonia* and *Mycoplasma genitalium* influence fibronectin binding. *Infection and Immunity*. 80(77): 3533-3541.
- Berman H., Westbrook J., Feng Z., Gilliland G., Bhat T., Weissig H., Shindyalov I. & Bourne P. 2000. The Protein Data Bank. *Nucleic Acids Research*. 1 (42): 28-235.
- Benson D., Karsch-Mizrachi I., Lipman D., Ostell J., Rao B. & Wheeler D. 2000. Genbank. *Nucleic Acids Research*. 1(8): 28-15.
- Barker W., Garavelli J., Huang H., McGarvey P., Orcutt B., Srinivasarao G., Xiao C., Yeh L., Ledley R., Janda J., Pfeiffer F., Mewes H., Tsugita A. & Wu C. 2000. The representation of protein complexes in the Protein Ontology (PRO). *Nucleic Acids Research*. 28(41): 12-371.
- Bon M. 2005. Guía de campo de los hongos de España y Europa, claves y descripción de 1500 especies y variedades de hongos. Editorial Omega. pp 19.
- Chen S. & Chen T. 2002. Retrieval of 3D protein structures. *Department of Electrical and Computer Engineering, Carnegie Mellon University*. 32(5): 933-936.
- Claus H. 2003. Lacasses and their occurrence in prokaryotes. *Archives Microbiology*. 4(6): 145-150.
- Colpa D., Fraaije M. & Bloois E. 2014. DyP-type peroxidases: a promising and versatile class of enzymes. *Journal of Industrial Microbiology and Biotechnology*. 41(1): 1-7.
- Chothia C. & Lesk A. 2003. The relation between the divergence of sequence and structure in proteins. *The Embo Journal*. 5(4): 823-6.

Corpet F. 1988. Múltiple sequence alignment with hierarchical clustering. *Nucleic Acids Research*. 16(22): 81-90.

Cordoba R. & Cultid G. 2015. Estudio comparativo de la actividad enzimática de lacasa (Lac), Lignina peroxidasa (Lip) y Manganese peroxidasa (MnP), de *Pleurotus ostreatus*, cultivado en residuos lignocelulósicos de raquis de palma de aceite, bagazo de fique y pulpa de café. 17(8): 71-36.

Diaz, R. 2009. Efecto del pH inicial de desarrollo de *Pleurotus ostreatus* en fermentación sumergida sobre su actividad de lacasas. Instituto Politecnico Nacional centro de investigación en biotecnología aplicada CIBA-IPN Tlaxcala. Tepetitla de Lardizábal Tlaxcala. 8: 22-24.

Diaz R., Téllez-Téllez M., Sánchez C., Bibbins-Martínez M., Díaz-Godínez G., & Soriano-Santos J. 2013. Influence of initial PH of the growing medium on the 12 activity, production and genes expression profiles of laccase of *Pleurotus ostreatus* in submerged fermentations. *Electric Journal of Biotechnology*. 16(4): 1-13.

Duran E. 2007. Evaluar la capacidad de remover el colorante azul índigo en aguas residuales, empleando una cepa nativa de *Pleurotus ostreatus*. Universidad Javeriana. Departamento de microbiología. pp 137.

Fraaije M. & Bloois E. 2013. DyP-type Peroxidases: A Promising and Versatile Class of Enzymes. Laboratory of Biochemistry, Groningen Biomolecular Sciences and Biotechnology Institute. University of Groningen. pp 41.

Fiser A., Do R. & Sali, A. 2000. Modeling of loops in protein structures. *Protein Science*. 9(9): 53-73.

Finn R., Clements J. & Eddy S. 2011. HMMER web server: interactive sequence similarity searching. *Nucleic Acids Research*. 39(4): 29-37.

Fischer D. & Eisenberg D. 1996. Protein fold recognition using sequence-derived predictions. *Protein Science*. 5(5): 947-55.

Guex N. & Peitsch M. 1997. SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis*. 18 (15): 14-23.

Greer J. 1990. Comparative modeling methods: application to the family of the mammalian serine proteases. *Proteins*. 7(4): 317-34.

Gianfreda L., Xu F. & Bollag J. 1999. Laccases: a useful group of oxidoreductive enzymes. *Bioremediation Journal*. 17(3): 11-25

Gasteiger E., Jung E. & Bairoch A. 2001. SWISS-PROT: connecting biomolecular knowledge via a protein database. *Current Issues in Molecular Biology*. 3(3): 47-55.

- Gómez A., Cedano J., Espadaler J., Hermoso A., Piñol J. & Querol E. 2008. Prediction of protein function improving sequence remote alignment search by a fuzzy logic algorithm. *The Protein Journal*. 27(2): 130-139.
- Hadar Y., Kerem Z., Gorodecki B. & Ardon O. 2012. Utilization of lignocellulosic waste by the edible mushroom, *Pleurotus*. *Biodegradation*. 24(3): 189-205.
- Haines E. 2006. An introductory tour of interactive rendering. *Computer Graphics and Applications*, Institute Electrical and Electronics Engineers. 26(1): 76-87.
- Hernández S., Ferragut G., Amela I., Cedano J., Perez-Pons J., Piñol J., Mozo-Villarias A. & Querol E. 2014. MultitaskProtDB: a database of multitasking proteins. *Nucleic Acids Research*. 42: 517-520.
- Hoofst R., Vriend G., Sander C. & Abola E. 1996. Errors in protein structures. *Nature*. 23(381): 272.
- Higuchi T. 2004. Microbial degradation of lignin: Role of lignin peroxidase, manganese peroxidase, and laccase. *Proceedings of the Japan Academy, series B. Physical and Biological Sciences*. 80(5): 204-214.
- Heinfling A., Martínez M., Martínez A., Bergbauer M. & Szewzyk U. 1998. Transformation of industrial dyes by manganese peroxidases from *Bjerkandera adusta* and *Pleurotus eryngii* in a manganese-independent reaction. *Applied Environmental Microbiology*. 64(8): 88-93.
- Hofrichter M., Ullrich R., Pecyna M., Liers L. & Lundell T. 2010. New and classic families of secreted fungal heme peroxidases. *Applied Microbiology and Biotechnology*. 87(3): 871-97.
- Hofrichter M., Lundell T. & Hatakka A. 2001. Conversion of milled pine wood by manganese peroxidase from *phlebia radiata*. *Applied Environmental Microbiology*. 84(1): 88-93.
- Jeanmougin F., Thompson J., Gouy M., Gibson D. & Higgins T. 1998. Multiple sequence alignment with Clustal X. *Trends in Biochemical Sciences*. 23(10): 403-405.
- Jones D. & Thornton J. 1993. *Biomolecular Structure and Modelling Unit, Department of Biochemistry and Molecular Biology, University College, London*. pp 68.
- Johnson M., Srinivasan N., Sowdhamini R. & Blundell T. 2003. Knowledge-based protein modeling. *Critical Reviews in Biochemistry and Molecular Biology*. 29(1): 1-68.
- Khan I., Chitale M., Rayon C. & Kihara D. 2012. Evaluation of function predictions by PFP, ESG, and PSI-BLAST for moonlighting proteins. *BioMed Central*. 6(7): 23-30

- Kirk T. & Cullen D. 1998. Enzymology and molecular genetic of wood degradation by white-rot fungi. En environmentally Friendly Technologies for pulp and paper Industry. 46: 273-307.
- Kirk T. & Farrell R. 1987. Enzymatic combustion: the microbial degradation of lignin. Annual Review Microbiology. (41): 465-505.
- Kreil D. & Etzold T. 2003. DATABANKS - a catalogue database of molecular biology databases. Trends in Biochemical Sciences. 24(4): 155-157.
- Kim S. & Shoda M. 1999. Purification and characterization of a novel peroxidase from *Geotrichum candidum* Dec 1 involved in decolorization of dyes. Applied Environmental Microbiology. 65(3): 29-35.
- Lathrop R., Temples F. & Smith J. 1995. A Bayes-optimal sequence-structure theory that unifies protein sequence-structure recognition and alignment. Bulletin of Mathematical Biology. 60(6): 39-71.
- Lemer C., Rooman M. & Wodak S. 2008. Protein structure prediction by threading methods. Evaluation of current techniques. The Proteins Journal. 23(3): 337-55.
- Liers C., Bobeth C., Pecyna M., Ullrich R. & Hofrichter M. 2010. DyP-like peroxidases of the jelly fungus *Auricularia auricula-judae* oxidize nonphenolic lignin model compounds and high-redox potential dyes. Applied Microbiology Biotechnology. 85(9): 69-79.
- Levitt M. 2003. Accurate modeling of protein conformation by automatic segment matching. Journal of Molecular Biology. 226(20): 507-533.
- Laskowski R., MacArthur M., Moss D. & Thornton J. 1993. PROCHECK: a program to check the stereochemical quality of protein structures. Journal of Applied Crystallography. 27(2): 269-305.
- Mester T. & Field J. 1998. Characterization of a novel manganese peroxidase-lignin peroxidase hybrid isozyme produced by *Bjerkandera* species strain in the absence of manganese. Journal of Biological Chemistry. 73(25): 419-424.
- Martínez A. 2002. Molecular biology and structure-function of lignin-degrading heme peroxidases. Enzyme and Microbial Technology. 30: 425-444.
- Niladevi K. 2009. Ligninolytic enzymes. Biotechnology for agro-industries, utilisation dordrecht. Springer Netherlands. 18: 541-545.
- Necochea R. & Canul J. 2004. Métodos fisicoquímicos en biotecnología, secuenciación de ácidos nucleicos. Instituto de Biotecnología. UNAM. 17: 8-48.
- Orengo C., Pearlm F., Bray J., Todd A., Martin A., Conte L. & Thornton J. 1999. Modern Protein Chemistry: Practical Aspects. Nucleic Acids Research. 27: 27-39.

- Pontius J., Richelle J. & Wodak S. 2003. Deviations from standard atomic volumes as a quality measure for protein crystal structures. *Journal of Molecular Biology*. 264(1): 121-136.
- Pearson W. 1990. Rapid and sensitive sequence comparison with FASTP and FASTA, methods in enzymology. *Nucleic Acids Research*. 30(4): 183-63.
- Peitsch M. 1997. Practical Application of Computer-Aided Drug Design. *Journal of Medicinal Chemistry*. 60: 15-84.
- Peitsch M. & Jongeneel C. 1993. A 3-D model for the CD40 ligand predicts that it is a compact trimer similar to the tumor necrosis factors. *International Immunology*. 5(1): 233-238.
- Rodríguez E. 2006. Caracterización molecular de lacasas de *Pleurotus eryngii*, expresión heteróloga de estas enzimas y aplicaciones en la degradación de contaminantes aromáticas. Tesis doctoral. Universidad Complutense de Madrid. Facultad de Ciencias Biológicas, Madrid. pp 25.
- Rost B. & Sander C. 1996. Biophys. Biomol. Struct. Bridging the protein sequence-structure gap by structure predictions. *Annual Reviews*. 349: 25-113.
- Ramachandran G. & Sasisekharan V. 2003. Conformation of polypeptides and proteins. *Advances in Protein Chemistry*. 54(23): 283-437.
- Rychlewski L., Zhang B. & Godzik A. 1998. Fold and function predictions for *Mycoplasma genitalium* proteins. *Folding and Design*. 3(85): 229-38.
- Santos O. 2000. Tese de Doutorado, Instituto Militar de Engenharia, Rio de Janeiro, Brasil. Modelagem Molecular da Di-hidrofolato Redutase do *Plasmodium falciparum* e Estudo de seus Inibidores pelo Método QSAR-3D. pp 112.
- Sanger F., Nicklen S. & Coulson A. 1977. DNA sequencing with chain terminating inhibitors. *Proceedings of the National Academy of Sciences*. 74(12): 5463-5467.
- Šali A. & Blundell T. 1993. Comparative protein modelling by satisfaction of spatial restraints. *Journal of Molecular Biology*. 234(3): 779-815.
- Sauder J., Arthur J. & Dunbrack, R. 2000. Large-scale comparison of protein sequence alignment algorithms with structure alignments. *The Protein Journal*. 40(1): 6-22.
- Santos O., Bicca R. & Figueroa J. 2001. Modelagem de proteínas por homologia. *Biophysical Chemistry*. 26(2): 91-305.
- Sánchez R. & Šali A. 1999. ModBase: a database of comparative protein structure models. *Bioinformatics*. 27(15): 60-61.
- Smith T. & Waterman M. 2003, Identification of common Molecular Subsequences. *Journal of Molecular Biology*. 147(1): 195-197.

- Sippl M. 1993. Recognition of errors in three-dimensional structures of proteins. *The Protein Journal*. 17(1): 355-362.
- Sugano Y. 2009. DyP-type peroxidases comprise a novel heme peroxidase family. *Cellular and Molecular Life Sciences*. 66(8): 87-93.
- Sugano Y., Matsushima Y., Tsuchiya K., Aoki H., Hirai M. & Shoda M. 2009. Degradation pathway of an anthraquinone dye catalyzed by a unique peroxidase DyP from *Thanatephorus cucumeris*. *Biodegradation*. 20(3): 433-440.
- Shah V. & Nerud F. 2002. Lignin degrading system of white-rot fungi and its exploitation for dye decolorization. *Canadian Journal of Microbiology*. 14(8): 857-870.
- Salvachúa D., Prieto A., Martínez Á. & Martínez M. 2013. Characterization of a novel DyP-type peroxidase from *Irpex lacteus* and its application in the enzymatic hydrolysis of wheat straw. *Applied and Environmental Microbiology*. 1(2): 164-177.
- Santos O. & Bica R. 2003. Modelagem de proteínas por homología. Departamento de Química Orgánica, Instituto de Química, Universidade Federal do Rio de Janeiro, Centro de Tecnologia. pp 26.
- Sánchez R. 2015. Hongos superiores como fuente de salud. Universidad Complutense, facultad de farmacia. pp 35.
- Trudy M. & Mckee J. 2003. La base molecular de la vida. Editorial McGRAW-HILL. pp 794.
- Varneo M., Quiroz M. & Álvarez C. 2010. Utilización de residuos forestales lignocelulosicos para producción del hongo ostra *Pleurotus ostreatus*. *Información tecnológica*. 21(2): 13-20.
- Vázquez E. & Contreras E. 2003. Cristalografía de rayos X, Química y Biología Molecular en Línea. Instituto de Química UNAM. <http://bq.unam.mx/evazquez>, (Consultado octubre 3, 2017).
- Vergara I., Navarro G. & Melo F. 2008. Obtención de un clasificador óptimo para la evaluación de la calidad de modelos tridimensionales de proteínas. 34(5): 20-188.
- Wang H., D Bell, & Murtagh F. 2008. Axiomatic Approach to Feature Subset Selection Based on Relevance. *Transactions on Pattern Analysis and Machine Intelligence*. 21(3): 271-276.

Zepeda O. 2013. Historia de la bioinformática, relación entre la biología y la informática, genómica funcional, biochips. Solo ciencia. <http://www.solociencia.com/biologia/bioinformatica-autor.htm> (Consultado: octubre 16, 2017).