


ARTÍCULO ORIGINAL

## RELACIÓN ESTRUCTURA–FUNCIÓN DE LAS PROTEÍNAS. USO DE LA TÉCNICA DE ELECTROFORESIS EN EL ESTUDIO DE SU EXPRESIÓN

### PROTEIN STRUCTURE-FUNCTION RELATIONSHIP. USE OF THE ELECTROPHORESIS TECHNIQUE IN THE STUDY OF THEIR EXPRESSION

Morela Fuchs Delgado<sup>1</sup> 

<sup>1</sup>Instituto Nacional de Investigaciones Agrícolas (INIA), Maracay, Estado Aragua, Venezuela.

#### Correspondencia

Morela Fuchs Delgado  
mfuchs@inia.gob.ve

**Para citar este artículo.** Fuchs, M. (2022). Relación estructura–función de las proteínas. Uso de la técnica de electroforesis en el estudio de su expresión. *Advances in Science and Innovation*, 1 (1),

#### RESUMEN

Las proteínas son componentes del metabolismo celular presentes en todos los organismos vivos, donde realizan importantes funciones que están relacionadas con su estructura. Estos compuestos han sido estudiados mediante diferentes técnicas y métodos, pero en los últimos años con el uso de nuevas técnicas, se ha ido cambiando el paradigma sobre estructura-función. En este estudio se realizó un análisis de los nuevos conocimientos sobre la estructura de las proteínas: los plegamientos, motivos y dominios que explican parte de la funcionalidad, las proteínas intrínsecamente desordenadas, también llamadas PINes o PIDs y las proteínas multifuncionales o Moonlighting, un concepto más flexible en el cual la estructura se adapta a las funciones que realiza. Se revisó el aporte de las nuevas técnicas bioinformáticas en el análisis y organización de una gran cantidad de información que se va produciendo con el estudio de la estructura, mediante cristalografía de rayos X y resonancia magnética nuclear (RMN). También, se analizó el estudio de la expresión de las proteínas a través de la técnica de electroforesis, que continúa siendo una herramienta valiosa y complementaria, para comprender el efecto de los factores ambientales sobre el desarrollo de los organismos.

**Palabras clave:** bioinformática, electroforesis, estructura, función, polipéptidos

#### ABSTRACT

Proteins are components of cellular metabolism present in all living organisms, where they perform important functions that are related to their structure. These compounds have been

studied using different techniques and methods, but in recent years with the use of new techniques, the structure-function paradigm has been changing. In this study, an analysis of the new knowledge about protein structure was carried out: the folds, motifs and domains that explain part of the functionality, intrinsically disordered proteins, also called PINEs or PIDs, and multifunctional proteins or Moonlighting, a concept more flexible in which the structure adapts to the functions it performs. The contribution of the new bioinformatic techniques in the analysis and organization of a large amount of information that is produced with the study of the structure, through X-ray crystallography and nuclear magnetic resonance (NMR), was reviewed. Also, the study of protein expression was analyzed through the electrophoresis technique, which continues to be a valuable and complementary tool to understand the effect of environmental factors on the development of organisms.

**Keywords:** bioinformatics, electrophoresis, function, polypeptides, structure

## INTRODUCCIÓN

Las proteínas son el producto final del proceso de decodificación de la molécula de ADN. La información contenida en la secuencia genética de tres bases nitrogenadas se traduce a aminoácidos y mediante un proceso de unión, a través de enlaces químicos se forma la estructura de una proteína específica, que va a realizar una o varias funciones en el metabolismo celular. Las proteínas realizan una amplia variedad de funciones en el organismo, incluida la replicación del ADN, el transporte de moléculas, la catalización de reacciones metabólicas y el suministro de soporte estructural a las células (Corral, 2017).

Esto pareciera ser un mecanismo simple que se realiza continuamente en el ciclo celular y en cada uno de los procesos vitales, pero aun cuando la estructura fundamental de las proteínas, los aminoácidos son un número limitado, la combinación de estos pueden dar origen a muchas proteínas y a proteínas que pueden cambiar su conformación estructural, para realizar múltiples funciones. El esquema rígido de la estructura de las proteínas ha cambiado hacia una visión

multifuncional, donde la estructura se modifica para actuar en determinados compartimentos celulares unida a otros compuestos biológicos. Hasta la fecha, se han utilizado varios términos no excluyentes para describir fenómenos recurrentes en los que las proteínas desempeñan dos o más funciones, especialmente pleiotrópicas, proteínas multidominio, promiscuidad y pluriempleo (Espinoza et al., 2020).

Los aminoácidos proteínogénicos, es decir los que forman proteínas, son un conjunto de veinte alfa-aminoácidos, cuya estructura difiere únicamente en el sustituyente (R) unido al carbono de la molécula.

Estos aminoácidos se unen en cadenas y pueden de acuerdo con su función, tener una estructura lineal o más compleja (Ayon, 2020).

Las funciones de las proteínas están determinadas por su estructura química y la configuración espacial permite la interacción con los otros componentes celulares, ya sea formando parte de la estructura, proteínas estructurales, o interviniendo en las

reacciones bioquímicas, como las enzimas (Kessel & Ben-Tal, 2018).

Para estudiar las proteínas se han utilizados diferentes técnicas que han permitido el avance en el conocimiento de su estructura y funciones en diferentes condiciones, durante el desarrollo de los organismos vivos. Esta comprensión ha sido de gran valor para entender la vida y las implicaciones de factores externos e internos en el equilibrio celular (Vega & Reyes, 2020).

### Estructura de las proteínas

La estructura completa de una proteína se puede describir en cuatro niveles diferentes de complejidad: estructura primaria, secundaria, terciaria y cuaternaria.

**Estructura primaria:** se define como la secuencia lineal de aminoácidos de la cadena polipeptídica de una proteína. Los aminoácidos están unidos por enlaces peptídicos. Los dos extremos de cada cadena polipeptídica se conocen como terminal amino (terminal N) y terminal carboxilo (terminal C) (Sanvitore & Farci, 2022).

**Estructura secundaria:** es la conformación espacial local del esqueleto polipeptídico excluyendo las cadenas laterales. Las estructuras secundarias regulares (también conocidas como elementos de estructura secundaria) comunes a muchas proteínas incluyen hélices  $\alpha$ , láminas  $\beta$  y giros. Debido a que todos los aminoácidos, excepto la glicina, son L-aminoácidos asimétricos, la cadena peptídica tiende a asumir una forma helicoidal asimétrica; algunas de las proteínas fibrosas consisten en hélices alargadas alrededor de un eje de tornillo recto. Tales características estructurales resultan de propiedades comunes a todas las cadenas peptídicas (Ealy, 2022), que se forman mediante los enlaces de hidrógeno que estabilizan el esqueleto polipeptídico (Sanvitore & Farci, 2022). Nekrasov et al. (2021), analizaron la estructura

de las proteínas globulares utilizando diagramas de empaquetamiento, demostrando cómo la disposición de segmentos de estructura secundaria a lo largo de la secuencia se relaciona con propiedades tridimensionales, como las hojas paralelas y antiparalelas.

**Estructura terciaria:** se refiere a la disposición tridimensional de todos los átomos que constituyen una molécula de proteína. Relaciona la coordinación espacial precisa de los elementos de la estructura secundaria y la ubicación de todos los grupos funcionales de una sola cadena polipeptídica. La mayor parte del conocimiento sobre la estructura secundaria y terciaria de las proteínas globulares se ha obtenido mediante el examen de sus cristales mediante difracción de rayos X (Ealy, 2022).

A este nivel de organización la presencia de pliegues es una característica muy importante asociada a su funcionalidad. El plegamiento de proteínas es un proceso mediante el cual una cadena polipeptídica se pliega para convertirse en una proteína biológicamente activa en su estructura 3D nativa. Las proteínas plegadas se mantienen unidas por diversas interacciones moleculares, lo que hace que las proteínas se plieguen, se entiende por su termodinámica, pero cómo se pliegan, se explica por la cinética de plegamiento de proteínas (Roshni, 2021).

Se ha determinado que el número de pliegues únicos es mucho menor que el número total de estructuras de proteínas y esto se debe a que no solo las proteínas funcionalmente relacionadas, tienen estructuras terciarias similares, sino que incluso, las proteínas con funciones muy diferentes a menudo comparten los mismos pliegues terciarios. Como consecuencia, la conservación estructural en el nivel terciario es quizás más profunda que en el primario (Guillén, 2019).

La identificación del pliegue de una proteína se

ha convertido en una herramienta invaluable, ya que está relacionada a la función y puede permitir mapear regiones funcionalmente importantes en la secuencia de aminoácidos. Basados en métodos predictivos computacionales, Yang et al. (2019) proponen dos algoritmos: MV-fold y MT-fold/MV-fold, para el reconocimiento de pliegues. Las diferentes características de las proteínas analizadas fueron tratadas como diferentes vistas de estas, incluyendo la información evolutiva, la información de la estructura secundaria y las propiedades fisicoquímicas.

### Plegamientos, motivos y dominios

Debido a la complejidad estructural de las proteínas a nivel terciario, se han establecido unas unidades estructurales de organización dentro de las estructuras tridimensionales de las proteínas denominados pliegues, motivos y dominios. Cada una de estas estructuras tiene una complejidad mayor, es decir los pliegues se organizan en motivos y los motivos en dominios (Olamoyesan & Rodger, 2021).

Los plegamientos son cambios que se producen en la estructura terciaria de las proteínas, que son termodinámicamente estables. El tiempo que tardan las proteínas en plegarse es variable y depende del número de aminoácidos que la conforman (Olivares-Quiroz & García-Coli, 2004).

Un motivo está formado por un conjunto de residuos de aminoácidos conservados que son importantes para la función. Los motivos pueden combinarse para formar estructuras globulares compactas que se denominan dominios (National Human Genome Research Institute, 2022). Típicamente, un dominio conservado contiene uno o más motivos.

Wetlaufer (1973), propone por primera vez el concepto de dominio al que considera una unidad independiente que constituye una estructura tridimensional. Es un conjunto de hélices  $\alpha$ , hojas

$\beta$  y hebras de enlace que pueden plegarse en forma simultánea e independiente de las partes restantes de la proteína y que además realiza una función específica (Olivares-Quiroz et al., 2004). Las proteínas pequeñas pueden contener un solo dominio, mientras que las proteínas más grandes a menudo contienen múltiples.

Mediante las diferentes definiciones reportadas en literaturas, redefine el dominio como “un arreglo globular estable de una gran porción de una proteína que contiene a menudo varios motivos”. Involucra secuencias de aminoácidos más grandes, siendo el motivo una forma de estructura secundaria, con formas particulares y comunes de asociación de hélices  $\alpha$ , láminas  $\beta$  y giros entre sí. Los conceptos de motivos y dominios formarían parte de esta estructura (Yu et al., 2019).

Levitt & Chothia (1976), en base a la simple consideración de los motivos conectados entre sí, desarrollaron una taxonomía de estructuras de proteínas y clasificaron los dominios en tres grupos principales: los dominios  $\alpha$ , cuyo centro está constituido exclusivamente por hélices  $\alpha$ , los dominios  $\beta$ , cuyo núcleo está formado generalmente por hojas  $\beta$  antiparalelas que suelen estar empaquetadas las unas con las otras y los dominios de estructura  $\alpha/\beta$  que están compuestos por combinaciones de los motivos  $\beta\alpha\beta$ .

La identificación de la estructura jerárquica de la proteína es compleja y para su identificación se han utilizado diferentes técnicas, desde métodos matemáticos, como el utilizado por Nekrasov et al. (2021), basado en el pentapéptido como unidad de secuencias de proteínas, algoritmos matemáticos (Islam et al., 1995; Romant & López, 2012), identificación de dominios mediante determinación de estructuras cristalinas (Khosla et al., 2020), alineamientos de las secuencias de ortólogos (González, 2013) y técnicas computacionales para la identificación de residuos estructuralmente

conservados (Via & Helmer-Citterich, 2004).

La presencia de algunos dominios ha sido asociada a determinadas funciones de las proteínas como por ejemplo las de defensa. El dominio TIR (Toll Interleukin Receptor), está implicado tanto en el reconocimiento como en la transducción de señales de defensa mediante su dimerización y están presentes en todas las formas de vida celular. Se encuentran más comúnmente en las proteínas de señalización, como unidades responsables de la formación de complejos de proteínas dependientes de la señal que permiten la amplificación y la propagación espacial de la señal (Khosla et al., 2020, Toshchakov & Neuwald, 2020). Mediante estudios funcionales de estos dominios se puede inferir el papel de estas secuencias TIR en diferentes procesos de defensa.

En un esfuerzo por comprender mejor la función del dominio TIR, se investigó el dominio TIR de la proteína RPV1 TIR-NLR de *Muscadina rotundifolia*, que confiere resistencia a *Plasmopara vitícola*, el agente causal del mildiú veloso en las vides cultivadas. *M. rotundifolia* es una especie de uva silvestre de América del Norte estrechamente relacionada con la especie de vid cultivada *Vitis vinifera*. Los resultados confirmaron que el dominio TIR es la región mínima requerida para la señalización de muerte celular, como mecanismo de protección contra la infección (Williams et al., 2016).

González (2013), realizó un análisis para determinar los dominios conservados por medio de alineamientos de las secuencias de ortólogos para la proteína flavonoide 3´5´hidroxilasa, F3´5´H de *Petunia hybrida*, enzima flavonoide-3´,5´-hidroxilasa involucrada en la biosíntesis de delphinidina, responsable del color azul en las plantas. La enzima piruvato quinasa, una enzima glucolítica que juega un papel importante en la regulación del flujo de fructosa-1,6-bisfosfato al piruvato,

ha sido una de las más estudiadas. Contiene un dominio de unión de nucleótidos totalmente  $\beta$ , un dominio de unión de sustrato  $\alpha / \beta$  y un dominio regulador  $\alpha / \beta$ , conectados por varios enlaces polipeptídicos. Cada dominio de esta proteína se encuentra en diversos conjuntos de familias de proteínas (Cardona, 2020).

El dominio WHY, encontrado en varias familias de proteínas, en particular las proteínas HinI, LEA 8 y LEA 14, ha sido detectado en respuesta hipersensibles al estrés hídrico, a la desecación en plantas y en archeas (Mertens et al., 2018).

**Estructura cuaternaria:** La estructura cuaternaria de muchas proteínas, especialmente aquellas mayores de 100 kDa en masa, son oligómeros que consisten en más de una cadena polipeptídica con una disposición espacial precisa de las subunidades dentro de la proteína. La estructura cuaternaria de las proteínas se forma mediante la unión por enlaces débiles de varias proteínas con estructura terciaria para formar un complejo proteico. Cada una de estas cadenas polipeptídicas recibe el nombre de protómero. A este nivel de organización se forman estructuras de gran importancia biológica como los microtúbulos y microfilamentos, el colágeno, las cápsulas de los virus y los complejos enzimáticos (Cardona, 2020).

#### **Estructura intrínsecamente desordenada**

Otra forma de estructura que ha sido estudiada en los últimos años en múltiples proteínas es una estructura no muy definida, llamada por algunos autores como estructura desordenada. Esto ha desafiado el dogma que relaciona la función de una proteína con una estructura tridimensional definida (Bruley et al., 2022). Algunos autores la señalan como proteínas de quinto orden. Estas presentan una estructura flexible o un plegamiento al azar. Estas proteínas han sido denominadas PIDs o IDPs, del inglés Intrinsically Disordered Proteins) o “proteínas no estructuradas” (PINEs o

IUPs, del inglés Intrinsically Unstructured Proteins, (Olamoyesan & Rodger, 2021).

Las PIDs o PINEs son proteínas en las que abundan las regiones intrínsecamente desordenadas RID o IDR (del inglés Intrinsically Disordered Regions) y que, por tanto, carecen total o parcialmente de una estructura tridimensional bien definida (Cuevas & Covarrubias, 2011).

Una herramienta desarrollada por Bruley et al. (2022), llamado pyHCA, permite especificar la relación entre el orden, representado por estructuras secundarias regulares (ya sea que participen en el núcleo hidrofóbico de estructuras 3D bien plegadas o que se formen condicionalmente en regiones intrínsecamente desordenadas) y el desorden.

La diversidad de estas proteínas de difícil de descifrar a pesar de que está codificada en las secuencias de aminoácidos. En general, presentan una composición de aminoácidos particular, son proteínas de baja complejidad, poseen una alta proporción de aminoácidos cargados o polares (Gln, Ser, Pro, Glu, Lys, Arg) y en algunos casos tienen una alta abundancia de ciertos aminoácidos pequeños (Gly, Ala) (Cuevas & Covarrubias, 2011).

El grupo de las proteínas LEA (late embryogenesis abundant) componen el grupo más abundante y caracterizado de IDP; está compuesta por una colección diversa de proteínas multidominio y multifuncionales que se encuentran en los tres dominios del árbol de la vida, pero son particularmente comunes en las plantas (Mertens et al., 2018).

Las proteínas LEA del grupo II son fitomoléculas esenciales que se acumulan principalmente en las últimas fases del desarrollo de la semilla y como reacción a tensiones externas extremas en los tejidos vegetativos (Aziz, 2021) y han sido el grupo

más estudiado. Las proteínas LEA del grupo 4 (LEA4), no se han encontrado similitudes significativas entre los miembros de los diferentes grupos, una característica común en la mayoría de ellos es su alta hidrofilia y alto contenido de glicina (Mertens et al., 2018; Amara et al., 2014)

Las estructuras intrínsecamente desordenadas (IDP) tienen una importante función en la señalización y las interacciones moleculares, la regulación de numerosas vías, la protección de células y proteínas, y la homeostasis celular. Los IDP también desempeñan un papel en la construcción de macromoléculas como el ribosoma, en la organización de la cromatina, en el montaje y desmontaje de microfilamentos y microtúbulos, en el transporte a través del poro nuclear, en la unión y transporte de moléculas pequeñas, en el funcionamiento de chaperonas de proteínas y ARN, y como enlazadores “entrópicos” flexibles que separan los dominios de proteínas funcionales (Olamoyesan & Rodger, 2021).

### Relación, estructura y función

La estructura de la proteína determina la función que realiza en la célula. Numerosos estudios del análisis de esta estructura tridimensional han cambiado el concepto que se tenía de estructura-función, donde cada proteína tenía solo una estructura y esta le permitía desarrollar una función. En la actualidad, se conoce que la proteína puede tener múltiples funciones y que su estructura cambia para adaptarse a la condición donde va a interactuar (Guillén, 2017).

Jeffery (1999), define por primera vez estas proteínas como *Moonlighting*, haciendo referencia a su capacidad de realizar varios empleos o trabajos, excluyendo de este concepto las proteínas que son resultado de fusiones de genes, proteínas homólogas, pero no idénticas, variantes de empalme, proteínas cuyas modificaciones postraduccionales pueden variar y las proteínas que

tienen una única función pero que pueden operar en diferentes lugares o utilizar diferentes sustratos. Estas proteínas poseen una elevada flexibilidad estructural, la cual les permite adoptar estructuras diferentes y, por tanto, reconocer ligandos diversos conservando la especificidad en el reconocimiento de estos. La primera función recibe el nombre de canónica y la segunda la de *Moonlighting* y generalmente la primera no se relaciona con la segunda (Hernández, 2016).

Las proteínas del pluriempleo o *Moonlighting*, se encuentran en todo el árbol evolutivo: bacterias, arqueas, mamíferos, reptiles, pájaros, peces, gusanos, insectos, plantas, hongos, protozoos e incluso virus. Incluyen enzimas que sirven como receptores, citocinas secretadas, factores de transcripción, estabilizadores del ADN, componentes del citoesqueleto o subunidades del proteosoma (Jeffery, 2018). Las proteínas y regiones de proteínas intrínsecamente desordenadas (simplemente denominadas IDP) forman parte de la estructura de estas proteínas, como ejemplo de este tipo de proteínas, el supresor de tumores p53, un factor de transcripción con regiones ordenadas y desordenadas que sirve como centro en las redes de interacción proteína-proteína (Bruley et al., 2022, Follis et al., 2014).

Entre otras proteínas *moonlighting*, la enzima gliceraldehído 3-fosfato deshidrogenasa (GAPDH), es la proteína con más funciones descritas hasta los momentos, pero estas funciones son en diferentes organismos (Jeffery, 2018; Hernández, 2016).

Otras proteínas con esta condición son la aconitasa, ubiquitin, albumina sérica, enolasa, hsp90 $\alpha$ , hsp70 y HMGB1 (Hernández, 2016). Muchos virus también presentan este tipo de proteínas, especialmente aquellos que tienen muy pocos genes en su genoma. Existe una base de datos para proteínas *moonlighting*, creada por Hernández et al. (2014) y actualizada por Franco et al. (2018), que sirve

como banco de datos de los trabajos realizados con este tipo de proteínas. De MultitaskProtDB-II, se ha encontrado que un porcentaje significativo de proteínas *moonlighting* (78 %) están relacionadas con enfermedades humanas y que un 48 % de ellas son dianas, para fármacos actuales. Además, un 25 % de proteínas *moonlighting* de la base de datos actualizada están implicadas en la virulencia de microorganismos patógenos (Nájar, 2019; Franco et al., 2018).

Las proteínas *moonlighting* es otro nivel de comprensión de cómo funciona la complejidad celular. La unión de estas proteínas a moléculas como el ADN o ARN evidencia su importancia en estos procesos (Franco et al., 2018).

#### Técnicas para el estudio de las proteínas

Para conocer la estructura de una proteína se deben analizar cuáles son los aminoácidos que forman la cadena polipeptídica. Este trabajo se ha realizado con técnicas laboriosas como la degradación de Edman, el método más antiguo basado en la eliminación sucesiva de aminoácidos N-terminales por métodos químicos. La secuenciación de un péptido por esta técnica requiere muchas horas y la sensibilidad está en el rango de 2 a 5 pmol de un péptido purificado. La secuenciación de péptidos por espectrometría de masas ha reducido, de manera considerable, la obtención de la secuencia de los péptidos derivados de una proteína y ha aumentado, considerablemente, el número de secuencias reportadas (Vega & Reyes, 2020).

La secuenciación/identificación de proteínas sigue siendo indispensable para analizar las proteínas expresadas en una célula, identificar proteínas específicas y determinar modificaciones postraduccionales. Las proteínas de interés suelen estar disponibles en cantidades bajas de microgramos o incluso menos. La electroforesis en gel, seguida de la transferencia a una membrana de PVDF para la secuenciación N-terminal o la

digestión en gel para generar péptidos que se pueden separar mediante cromatografía líquida de alta presión (HPLC), permite el análisis posterior de las proteínas separadas. El análisis estructural se puede realizar mediante degradación de Edman o espectrometría de masas (MS) (Dams et al., 2019; Donnelly et al., 2019; EL Sharif et al., 2022).

Esta información ha sido complementada con la cristalografía de rayos X, la espectroscopía de RMN y la espectrometría de dicroísmo para el análisis de la estructura. El dicroísmo circular (CD) es un método de espectroscopia de absorción basado en la absorción diferencial de luz polarizada circularmente izquierda y derecha (ATA Scientific, 2019).

### Electroforesis de proteínas

Para realizar la identificación y caracterización de proteínas en un medio soluble, la técnica de electroforesis ha sido una herramienta muy útil en estudios que relacionan la proteína con la función, cuando se analizan la expresión de las proteínas de muestras obtenidas en diferentes condiciones bióticas y abióticas. Esta técnica se basa en la movilidad de las proteínas en un campo cargado eléctricamente. Cada proteína tiene una carga eléctrica que depende del balance entre las cargas positivas y negativas de la molécula, de los grupos amino ( $\text{NH}_4$ ) y carboxilo ( $\text{COOH}$ ). Las cadenas laterales de proteínas cargadas positivas y negativamente, hacen que se comporten como aminoácidos en un campo eléctrico; es decir, migran durante la electroforesis a valores bajos de pH al cátodo (terminal negativo) y a valores altos de pH al ánodo (terminal positivo) (Ealy, 2022).

Han sido muchos los trabajos realizados con la técnica de electroforesis, a partir de los trabajos metodológicos realizados por Laemmli (1970), con algunas modificaciones que se ajustan dependiendo del tipo de proteína y las características de la muestra a analizar (Navarro & Suárez, 2019., González &

Fillat, 2018; Moreno et al., 2017; Martínez et al., 2017).

La técnica de electroforesis comprende los siguientes procesos:

1. Extracción de las proteínas del órgano, tejido o célula a estudiar.
2. Preparación de la muestra, esta debe contener un buffer específico para el tipo de proteína, un colorante de corrida y glicerol o azúcar para incrementar la densidad.
3. Preparación del gel de corrida si la electroforesis es continua o de dos o más geles si es discontinua. El gel más utilizado en la separación de proteínas es la poliacrilamida.
4. Establecer las condiciones de corrida: voltaje, amperaje y tiempo.
5. Tinción de los geles (Stegeman et al., 1985)

La extracción de la muestra es una de las etapas fundamentales, debido a que la solubilidad y la estabilidad de las proteínas depende de diversos factores como la temperatura, la constante dieléctrica del medio, el pH, la fuerza iónica y la presencia de agentes reductores y la concentración de la proteína extraída es determinante para que pueda ser visualizada en los geles de electroforesis (Colás & Van Der Straeten, 2019; Martínez et al., 2017).

Empleando técnicas como las que aplican geles de poliacrilamida en presencia del detergente aniónico dodecilsulfato sódico (SDS-PAGE) (es el acrónimo en inglés de *sodium dodecyl sulfate polyacrylamide gel electrophoresis*), se pueden separar las proteínas dependiendo de su peso molecular. EL SDS actúa rompiendo enlaces no covalentes en las proteínas, desnaturalizándolas, es decir, que las moléculas proteicas pierden su conformación nativa. Se pueden separar las proteínas en base no solo a su carga eléctrica, sino también en base



a su peso y utilizando un patrón, conocer el peso molecular de cada una de las proteínas separadas. Para este tipo de análisis, la muestra de proteínas debe ser preparada con SDS-mercaptoetanol. Todo el sistema de electroforesis, es decir la muestra, el gel de corrida y el buffer de corrida debe contener SDS (Vega & Reyes, 2020; Matsumoto et al., 2019; Brunelle & Green 2014).

Cuando la preparación de las muestras se realiza para estudios enzimáticos, se debe mantener la estabilidad estructural de la proteína, para que esta pueda expresarse y visualizarse durante la tinción de los geles, mediante el sustrato específico para cada tipo de enzima. La técnica SDS-PAGE no puede ser utilizada en este tipo de estudio. El uso de la electroforesis de isoenzimas se incrementó, debido al desarrollo de métodos, que permitieron visualizar una gran cantidad de enzimas de interés. Utilizando esta metodología se evaluó el efecto de factores limitantes sobre el desarrollo del cultivo y la variabilidad genética en plantas *in vitro* y en colecciones de germoplasma (Castilla et al., 2014; Giraldo et al., 2011)

Las variaciones en la técnica de electroforesis le dan una mayor amplitud de uso, la electroforesis discontinua (PAGE-DISC), separación mediante el punto isoeléctrico de la proteína (IF), la electroforesis bidimensional y el análisis de cambio de movilidad electroforética (EMSA), son herramientas actualmente utilizadas en el estudio de las proteínas (Montaldo & Lugo, 2019).

El EMSA es un procedimiento bioquímico utilizado para estudiar la unión entre proteínas y ácidos nucleicos. Este ensayo se basa en la migración diferencial de complejos de ARN/proteína y ARN libre durante la electroforesis en gel nativo. Mediante el uso de una sonda de ARN radiomarcada, los complejos de ARN/proteína se pueden visualizar mediante autorradiografía. El EMSA se desarrolló originalmente para estudiar la unión de proteínas

al ADN con secuencias de ADN diana (Fried & Crothers 1981). Esta técnica generalmente usa geles de poliacrilamida (Seo et al., 2019), pero también han sido desarrollados protocolos más rápidos con geles de agarosa (Ream et al., 2016) o combinación con otras técnicas (Charlier & Bervoets, 2022).

Esta técnica ha sido ampliamente utilizada en el estudio de proteínas moonlighting cuya función está asociada a la unión con moléculas de ARN y ADN. Wegener y Dietz (2022), realizaron una revisión sobre la interacción entre las enzimas glucolíticas, su interacción con el ARN en la regulación postranscripcional y los métodos utilizados para su análisis.

La proteómica analiza las proteínas presentes en una muestra mediante separación electroforética. Los datos proteómicos proporcionan contenidos moleculares específicos para un sistema celular e incluyen interacciones en par de ADN-proteína, proteína-proteína y ligación proteína (sustrato-enzima), los cuales determinan la conectividad local, que existe entre las especies moleculares, facilitando una plataforma informática dentro del sistema celular (Kelly, 2020). La proteómica ha permitido estudiar la respuesta de las proteínas, a través de los cambios producidos por la interacción con factores bióticos y abióticos (Fuchs et al., 2021; Komatsu & Hossain, 2017).

En esta técnica, mediante electroforesis bidimensional, se separan en un primer gel, las proteínas en base a su peso molecular con SDS-PAGE y posteriormente, en otra corrida en base a su enfoque isoeléctrico (IEF, del inglés isoelectrofocusing), esto permite una mayor cantidad de spots o manchas que se analizan y que corresponderían a proteínas diferentes. Se analizan cada uno de los péptidos separados del gel, mediante digestión con espectrometría de masas. La técnica de MALDI-TOF MS (matrix assisted laser desorption ionization-time of flight-mass

spectrometry, en español: desorción/ionización láser asistida por una matriz con detección de masas por tiempo de vuelo), ha resultado eficiente para muchos estudios, donde se requiere rapidez y precisión para el análisis de gran cantidad de muestras (Bonifaz et al., 2019; Maldonado, 2018; Relloso et al., 2015). Esta gran cantidad de datos generados son posteriormente analizados en una base de datos de proteínas para determinar a qué proteínas corresponden.

### Bioinformática

Una gran cantidad de datos se están generando del análisis de los componentes biológicos incluyendo interacciones, redes y vías, para analizar los principios elementales del sistema celular. Las matemáticas y las ciencias computacionales son

herramientas que ayudan a comprender toda la interacción entre los compuestos biológicos, que son complejas.

Las estructuras de las proteínas que se han determinado a través de las metodologías experimentales, principalmente por cristalografía rayos X (XRC), espectroscopia de resonancia magnética nuclear (RMN) y microscopía electrónica, a menudo son enviados al banco de datos de proteínas (Protein Data Bank, 2022), que es el depósito más grande para estructuras de proteínas determinadas experimentalmente, enriqueciendo la información disponible que a su vez servirá de base para otros estudios, esta información está disponible por método y por tipo molecular (Tabla 1).

**TABLA 1**

Distribución de datos PDB por método experimental y tipo molecular

| Tipo molecular             | Radiografía | RMN   | EM   | Múltiples Métodos | Neutrón | Otros | Total  |
|----------------------------|-------------|-------|------|-------------------|---------|-------|--------|
| Proteína sola              | 146986      | 11965 | 7501 | 186               | 72      | 32    | 166742 |
| Proteína/<br>oligosacarido | 8681        | 31    | 1313 | 5                 | 0       | 0     | 10030  |
| Proteína/NA                | 7756        | 277   | 2376 | 3                 | 0       | 0     | 10412  |
| Ácidos nucleicos           | 2449        | 1414  | 62   | 11                | 2       | 1     | 3939   |
| Otros                      | 154         | 31    | 5    | 0                 | 0       | 0     | 190    |

**Nota:** Protein Data Bank (2022)

Sin embargo, a pesar del rápido aumento de estructuras caracterizadas por métodos experimentales, todavía existe una brecha entre la secuencia de proteínas, (entradas de proteínas en UniProt a partir de enero de 2019) y la estructura (entradas de proteínas a PDB a partir de enero de 2019). Es aquí donde la bioinformática resulta ser una herramienta prometedora para el análisis de la información resguardada en las bases de datos. Los

métodos computacionales permiten la predicción de la estructura y función de las proteínas mediante el modelado comparativo, que consiste en determinar similitudes entre la secuencia objetivo y al menos una estructura conocida (Castro et al., 2019).

El diseño in silico, es una herramienta que permite caracterizar la estructura de la proteína, utilizando

la bioinformática, ya que por otras vías sería más laborioso (Navarro & Suárez, 2019; Jiménez & Chaparro, 2016).

El término *in silico* se refiere a los distintos procesos que se realizan a través de ordenadores computacionales, gracias a programas y bases de datos. Este tipo de análisis tiene base en la bioinformática, cuenta con diversas aplicaciones, de las cuales, tres son muy importantes: el modelado, la simulación y la visualización.

Actualmente, hay una serie de servidores que permiten la predicción de motivos, patrones y dominios en la secuencia, algunos de estos son: PROSITE, PRINTS, InterPro, ProDom y Pfam (Castro et al., 2019). Sklepari et al. (2016), combinaron cuatro técnicas bien establecidas: dispersión dinámica de la luz (DLS), dicroísmo circular (CD), espectroscopia de resonancia magnética nuclear (RMN) y modelado molecular, para proporcionar información sobre la proteína insulina.

El Banco de Datos de Dicroísmo Circular de Proteínas (PCDDDB), es otro recurso establecido para las comunidades de biología, biofísica, química, bioinformática y biología molecular. Los datos incluidos en el PCDDDB complementan los resultados de cristal, microscopía crioelectrónica, espectroscopía de RMN, caracterizaciones y clasificaciones bioinformáticas y otra información estructural disponible para las proteínas a través de enlaces a otras bases de datos (Gomes et al., 2022)

Las comparaciones de los resultados obtenidos utilizando las técnicas establecidas para el estudio de la estructura de las proteínas asociadas a la función o funciones, bajo determinadas condiciones experimentales, con la información disponible en las bases de datos ha resultado en un avance en el conocimiento de estos compuestos en el metabolismo celular.

## CONCLUSIONES

- El concepto de las proteínas como componentes esenciales del metabolismo celular ha variado debido a los últimos avances en el estudio de su estructura relacionada con su función, por lo que ha dejado de ser rígido a más flexible, adaptado a diferentes condiciones dentro del metabolismo celular.
- Las proteínas *moonlighting* también llamadas multifuncionales ha sido corroborado por estudios experimentales, presentes en los diferentes niveles de organización de la vida.
- La existencia en la estructura de motivos y dominios específicos para funciones determinadas ha permitido comparar las funciones bajo condiciones bióticas y abióticas que pueden ser limitantes al desarrollo de los organismos.
- Las técnicas bioinformáticas han ayudado para organizar la gran cantidad de información que ha sido generada con las técnicas convencionales de análisis de la estructura proteica y las bases de datos se han enriquecido con esta información que retroalimenta los estudios de las proteínas.

## REFERENCIAS BIBLIOGRÁFICAS

- Amara, I., Zaidi, I., Masmoudi, K., Ludevid, M., Pagès, M., Goday, A., & Brini, F. (2014) Insights into Late Embryogenesis Abundant (LEA) Proteins in Plants: From Estructura a las Funciones. *Revista Americana de Ciencias Vegetales*, 5, 3440 - 3455. DOI: 10.4236/ajps.2014.522360.
- ATA Scientific (2019). *Protein analysis techniques explained. Biomolecular Science guide.* <https://www.atascientific.com.au/3-protein-analysis-techniques/>
- Aziz, A., Sabeem, M., Mullath, S., Brini, F., & Masmoudi, K. (2021). Plant Group II LEA Proteins: Intrinsically Disordered Structure

- for Multiple Functions in Response to Environmental Stresses. *Biomolecules*, *11* (11), 1662. DOI: 10.3390/biom11111662.
- Bonifaz, A., Montelongo-Martínez F., Araiza, J., González, G., Treviño-Rangel R., Flores-Garduño A., Camacho-Cruz, A., & Tirado-Sánchez, A. (2019). Evaluación de MALDI-TOF MS para la identificación de levaduras patógenas oportunistas de muestras clínicas. *Revista Chilena de Infectología*, *36* (6), 790 - 793. <https://dx.doi.org/10.4067/S0716-10182019000600790>
- Brunelle J., & Green, R. (2014). One-dimensional SDS-polyacrylamide gel electrophoresis (1D SDS-PAGE). *Methods Enzymol*, *541*, 151 - 159. DOI: 10.1016/B978-0-12-420119-4.00012-4.
- Cardona, F. (2020). Las proteínas. De la estructura primaria a la cuaternaria. Aplicaciones. Escuela Técnica Superior de Ingeniería Agronómica y del Medio Natural. Universitat Politècnica de València. <http://hdl.handle.net/10251/147139>
- Castilla, Y., González M., & Lara, R. (2014). Determinación de estabilidad genética en vitroplantas de clavel español (*Dianthus caryophyllus* L.), micropropagadas con Biobras-16. *Cultivos Tropicales*, *35* (1), 67 - 74. [http://scielo.sld.cu/scielo.php?script=sci\\_arttext&pid=S0258-59362014000100010&lng=es&tlng=es](http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S0258-59362014000100010&lng=es&tlng=es).
- Castro, J., Maddox, J., Dylan, E., Segundo, L., Rodríguez, H., Casuso, M., Paredes, J., & Cobos, M. (2019). Caracterización *in silico* y análisis de la expresión de la subunidad alfa de la acetilcoenzima a carboxilasa heteromérica de dos microalgas. *Acta Biológica Colombiana*, *24* (2), 275 - 290. <https://doi.org/10.15446/abc.v24n2.74727>
- Charlier, D. & Bervoets, I. (2022). Separation and Characterization of Protein–DNA Complexes by EMSA and In-Gel Footprinting. In: Peeters, E., Bervoets, I (eds) Prokaryotic Gene Regulation. *Methods in Molecular Biology*, 2516. Humana, New York, NY. [https://doi.org/10.1007/978-1-0716-2413-5\\_11](https://doi.org/10.1007/978-1-0716-2413-5_11)
- Colás A., & Van Der Straeten, D. (2017). Optimization of non-denaturing protein extraction conditions for plant PPR proteins. *PLoS One*, *12* (11). DOI: 10.1371/journal.pone.0187753.
- Corral, R. (2017). Modelo de espacio vectorial para la representación y clasificación de las estructuras de proteínas. [Tesis doctoral. Universidad Nacional Autónoma de México]. <https://www.passeidireto.com/archivo/111492383/modelo-de-espacio-vectorial-para-la-representacion-y-clasificacion-de-estructura/6>
- Cuevas, C., & Covarrubias, A. (2011). Las proteínas desordenadas y su función: una nueva forma de ver la estructura de las proteínas y la respuesta de las plantas al estrés. *Revista Especializada en Ciencias Químico-Biológicas*, *14* (2), 97 - 105. <https://www.medigraphic.com/cgi-bin/new/resumen.cgi?IDARTICULO=36540>
- Dams, M., Dores-Sousa, J. L., Lamers, R.-J., Treumann, A., & Eeltink, S. (2019). High-Resolution Nano-Liquid Chromatography with Tandem Mass Spectrometric Detection for the Bottom-Up Analysis of Complex Proteomic Samples. *Chromatographia*, *82* (1), 101 – 110. <https://doi.org/10.1007/s10337-018-3647-5>
- Donnelly, D., Rawlins, C., DeHart, C., Fornelli, L., Schachner, L., Lin, Z., Lippens, J., Aluri, K., Sarin, R., Chen, B., Lantz, C., Jung, W., Johnson, K., Koller, A., Wolff, J., Campuzano, I., Auclair, J., Ivanov, A., Whitelegge, J., Agar, J. (2019). Best practices and benchmarks for intact protein analysis for top-down mass spectrometry. *Nature Methods*, *16* (7), 587 - 594. <https://doi.org/10.1038/s41592-019-0587-5>

org/10.1038/s41592-019-0457-0

- Ealy, J. (2022). *The Artistic and Scientific Nature of Protein Structure: A Historical Overview*, 625 – 648. [https://doi.org/10.1007/978-3-030-94651-7\\_28](https://doi.org/10.1007/978-3-030-94651-7_28)
- Espinosa-Cantú, A., Cruz, E., Noda-Garcia, L., & De Luna, A. (2020). Multiple Forms of Multifunctional Proteins in Health and Disease. *Front Cell Dev Biol*, 10 (8), 451. DOI: 10.3389/fcell.2020.00451.
- EL Sharif, H., Giosia, F., & Reddy, S. (2022). Investigation of polyacrylamide hydrogel-based molecularly imprinted polymers using protein gel electrophoresis. *Journal of Molecular Recognition*, 35 (1). <https://doi.org/10.1002/jmr.2942>
- Follis A., Llambi, F., Ou, L., Baran, K., Green, D., Kriwacki, R. (2014). The DNA-binding domain mediates both nuclear and cytosolic functions of p53. *Nature Structural & Molecular Biology*, 21 (6), 535 - 43. DOI: 10.1038/nsmb.2829.
- Franco, L., Hernandez, S., Calvo, A., Severi, M., Ferragut, G., Perez, J., Pinol, J., Pich, O., Mozo, A., Amela, I., Quero, E., & Cedano, J. (2018). MultitaskProtDB-II: an update of a database of multitasking/moonlighting proteins. *Nucleic Acids Research*, 46 (D1), D645 - D648. DOI: 10.1093/nar/gkx1066.
- Fried, M., & Crothers, D. (1981). Equilibria and kinetics of lac repressor-operator interactions by polyacrylamide gel electrophoresis. *Nucleic Acids Research*, 9 (23), 6505 - 6525. DOI: 10.1093/nar/9.23.6505.
- Fuchs, M., Almeida, I., & Fernández, M. (2021). Screening of rice proteins using 2-d gels after inoculation with *Rhizotonia solani*, *Revista Tayacaja*, 4 (1), 100 - 113. <https://doi.org/10.46908/tayacaja.v4i1.156>
- Giraldo, M., Ligarreto, G., Cayón, G., & Melo, C. (2011). Análisis de la carga genética de la colección colombiana de musáceas usando marcadores isoenzimáticos. *Acta Agronómica*, 60 (2), 108 - 119. <https://agris.fao.org/agris-search/search.do?recordID=CO2021A00891>
- Gomes, S., Miles, A., Janes, R., Wallace, B. (2022). The PCDDDB (Protein Circular Dichroism Data Bank): A Bioinformatics Resource for Protein Characterisations and Methods Development, *Journal of Molecular Biology*, 434 (11). <https://doi.org/10.1016/j.jmb.2022.167441>.
- González, A. (2013). Análisis *in silico* de los posibles dominios conservados y de regulación de la proteína flavonoidE-3',5'-hidroxilasa (F3'5'H) en *Petunia híbrida*. [Tesis de Maestría. Centro de Investigación y Asistencia en Tecnología y Diseño del estado de Jalisco]. <https://ciatej.repositorioinstitucional.mx/jspui/bitstream/1023/488/1/Adriana%20Gonz%C3%A1lez%20Dur%C3%A1n.pdf>
- González, A., & F. Fillat, (2018). Aspectos metodológicos de la expresión de proteínas recombinantes en *Escherichia coli*. *Revista de Educación Bioquímica (REB)*, 37 (1), 14 - 2. <https://www.medigraphic.com/pdfs/revedubio/reb-2018/reb181c.pdf>
- Guillén, M. (2017). Estructura y propiedades de las proteínas. [https://www.uv.es/tunon/pdf\\_doc/proteinas\\_09.pdf](https://www.uv.es/tunon/pdf_doc/proteinas_09.pdf)
- Hernández, S. (2016). Análisis bioinformáticos de las proteínas multifuncionales. [Tesis Doctoral. Universitat Autònoma de Barcelona]. <https://www.tesisenred.net/handle/10803/382811#page=1>
- Hernández, S., Ferragut, G., Amela, I., Perez-Pons, J., Pinol, J., Mozo-Villarias, A., Cedano, J., & Querol, E. (2014). MultitaskProtDB: a database

- of Proteínas multitarea. *Núcleo Ácidos Res.*, *42*, 517 - 520. DOI: 10.1093/nar/gkt1153.
- Islam S., Luo J. & Sternberg M. (1995). Identification and analysis of domains in proteins, *Protein Engineering, Design and Selection*, *8* (6), 513 – 526. <https://doi.org/10.1093/protein/8.6.513>
- Ayon, N. (2020). Features, roles and chiral analyses of proteinogenic amino acids. *AIMS Molecular Science*, *7* (3), 229 – 268. <https://doi.org/10.3934/molsci.2020011>
- Jeffery, C. (2018). Protein moonlighting: what is it, and why is it important? *Philos Trans R Soc Lond B Biol Sci.*, *373* (1738), 20160523. DOI: 10.1098/rstb.2016.0523.
- Jeffery, C. (1999). Moonlighting Proteins *Trends Biochemical Sciences*, *24* (1), 8 - 11. DOI:10.1016/S0968-0004(98)01335-8
- Jiménez, J., & Chaparro-Giraldo, A. (2016). Diseño *in silico* y evaluación funcional de genes semisintéticos que confieran tolerancia a fosfotricina. *Revista Colombiana de Biotecnología*, *18* (2), 90 - 96. <https://www.redalyc.org/articulo.oa?id=77649147011>
- Karasawa, M., Vencovsky, R., Silva, C., Cardim, D., Bressan, E., Oliveira, G., & Veasey, E. (2012). Comparison of microsatellites and isozymes in genetic diversity studies of *Oryza glumaepatula* (Poaceae) populations. *Revista de Biología Tropical*, *60* (4), 1463 - 78. DOI: 10.15517/rbt.v60i4.2055.
- Kelly, R. (2020). Single-cell Proteomics: Progress and Prospects. *Molecular & Cellular Proteomics*, *19* (11), 1739 – 1748. <https://doi.org/10.1074/mcp.R120.002234>
- Kessel, A., & Ben-Tal, N. (2018). Introduction to Proteins. Chapman and Hall/CRC. <https://doi.org/10.1201/9781315113876>
- Khosla, A., Morffy, N., Li, Q., Faure, L., Chang, S., Yao, J., Zheng, J., Cai, M., Stanga, J., Flematti, G., Waters, M., & Nelson, D. (2020). Structure–Function Analysis of SMAX1 Reveals Domains That Mediate Its Karrikin-Induced Proteolysis and Interaction with the Receptor KAI2. *The Plant Cell*, *32* (8), 2639 – 2659. <https://doi.org/10.1105/tpc.19.00752>
- Komatsu, S., & Hossain, Z. (2017). Preface—Plant Proteomic Research. *International Journal of Molecular Sciences*, *18* (1), 88. <http://dx.doi.org/10.3390/ijms18010088>
- Laemmli, U. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*, *227*, 680 - 685. <https://www.nature.com/articles/227680a0>
- Levitt, M., & Chothia, C. (1976). Structural patrons in globular protein. *Nature*. *261* (5561), 552 - 558, DOI: 10.1038 / 261552a0.
- Nájar, A. (2019). Análisis de la virulencia de proteínas multifuncionales mediante informática. [Tesis de Maestría, Universitat Oberta de Catalunya]. <http://openaccess.uoc.edu/webapps/o2/bitstream/10609/98607/7/anajarTFM0619memoria.pdf>
- National Human Genome Research Institute (2022). Protein. <https://www.genome.gov/es/genetics-glossary/Protein>
- Navarro, S., & Suárez, A. (2019). Evaluación *in silico* de la estructura y función de la proteína hipotética B7FQK1 de *Phaeodactylum tricornutum*. [Tesis de Maestría. Universidad Libre]. <https://repository.unilibre.edu.co/handle/10901/17781>
- Nekrasov, A., Kozmin, Y., Kozyrev, S., Ziganshin, R., Brevern, A., & Anashkina, A. (2021). Hierarchical Structure of Protein Sequence. *International Journal of Molecular Science*, *22* (15), 8339.

- <https://doi.org/10.3390/ijms22158339>
- Maldonado, N., Robledo, C. & Robledo, J. (2018). La espectrometría de masas MALDI-TOF en el laboratorio de microbiología clínica. *Infectio*, 22 (1), 35 - 45. <http://www.scielo.org.co/pdf/inf/v22n1/0123-9392-inf-22-01-00035.pdf>
- Martínez, A., Martínez, S., & Ardila, H. (2017). Condiciones para el análisis electroforético de proteínas apoplásticas de tallos y raíces de clavel (*Dianthus caryophyllus* L) para estudios proteómicos. *Revista Colombiana de Química*, 46 (2), 5 - 16 <https://pesquisa.bvsalud.org/portal/resource/pt/biblio-900819>
- Matsumoto, H., Haniu, H., & Komori, N. (2019). Determination of Protein Molecular Weights on SDS-PAGE. *Methods in Molecular Biology*, 1855, 101 - 105. DOI: 10.1007/978-1-4939-8793-1\_10.
- Mertens, J., Aliyu, H., & Cowan, D. (2018). LEA Proteins and the Evolution of the WHy Domain. *Applied and Environmental Microbiology*, 84 (15), e00539 - 18. DOI: 10.1128/AEM.00539-18.
- Montaldo, C., & Lugo, M. (2019). Electroforesis: fundamentos, avances y aplicaciones. *Epistemus*, 13 (26), 48 - 54. <https://doi.org/10.36790/epistemus.v13i26.96>
- Moreno, C., Fernández, R., & Valbuena, O. (2017). Caracterización electroforética de las proteínas del endospermo de variedades de arroz venezolanas. *Bioagro*, 29 (1), 37 - 44. [http://ve.scielo.org/scielo.php?script=sci\\_arttext&pid=S1316-33612017000100004&lng=es&tlng=es](http://ve.scielo.org/scielo.php?script=sci_arttext&pid=S1316-33612017000100004&lng=es&tlng=es)
- Olamoyesan, A., Ang, D., & Rodger, A. (2021). Circular dichroism for secondary structure determination of proteins with unfolded domains using a self-organising map algorithm SOMSpec. *RSC advances*, 11 (39), 23985 - 23991. <https://doi.org/10.1039/d1ra02898g>.
- Olivares-Quiroz, L., & García-Coli, L. (2004). Plegamiento de las proteínas: Un problema interdisciplinario. *Revista de la Sociedad Química de México*, 48(1), 95-105. <https://www.scienceopen.com/document?vid=d35e4a4d-8f1e-4b8c-84ba-2cf81ee4c101>
- Protein Data Bank (PDB). PDB Data Distribution by Experimental Method and Molecular Type <https://www.rcsb.org/stats/summary>
- Ream J., Lewis, L., & Lewis, K. (2016). Rapid agarose gel electrophoretic mobility shift assay for quantitating protein: RNA interactions. *Analytical Biochemistry*, 15 (511), 36 - 41. DOI: 10.1016/j.ab.2016.07.027
- Relloso, M., Nievas, J., Fares, T., Farquharsona, V., Mujica, M., Romano, V., Zaratea, M., & Smayevsky, J. (2015). Evaluación de la espectrometría de masas: MALDI-TOF MS para la identificación rápida y confiable de levaduras. *Revista Argentina de Microbiología*, 47 (2), 103 - 107. <https://doi.org/10.1016/j.ram.2015.02.004>
- Roshni, K. (2021). Protein folding, misfolding, and coping mechanism of cells—A short discussion. *Open Journal of Cell and Protein Science*, 4 (1), 001 - 004. DOI:10.17352/ojcps.000003
- Seo, M., Lei, L., & Egli, M. (2019). Label-Free Electrophoretic Mobility Shift Assay (EMSA) for Measuring Dissociation Constants of Protein-RNA Complexes. *Current Protocols in Nucleic Acid Chemistry*, 76 (1), e70. <https://doi.org/10.1002/cpnc.70>

- Sklepari, M., Rodger, A., Reason A., Jamshidi S., Prokes, I., & Blindauer, C. (2016). Biophysical characterization of a protein for structure comparison: methods for identifying insulin structural changes. *Analytical Métodos*, 8 (41), 7460 - 7471 DOI:10.1039/C6AY01573E
- Stegeman H., Burgermeister, W., Shah A., Franksen, H. & Krogerrecklenfort, E. (1985). Manual. PHOMA-PHOR
- Toshchakov, V., & Neuwald, A. (2020). A survey of TIR domain sequence and structure divergence. *Immunogenetics*. 72, 181 – 203. <https://doi.org/10.1007/s00251-020-01157-7>
- Vega, N., & Reyes, E. (2020). Introducción al análisis estructural de proteínas y glicoproteínas. Universidad Nacional de Colombia. [http://ciencias.bogota.unal.edu.co/fileadmin/Facultad\\_de\\_Ciencias/Publicaciones/Imagenes/Portadas\\_Libros/Quimica/Introduccion\\_al\\_analisis\\_estructural\\_de\\_proteinas\\_y\\_glicoproteinas/Analisis\\_estructural\\_proteinas\\_y\\_glicoproteinas.pdf](http://ciencias.bogota.unal.edu.co/fileadmin/Facultad_de_Ciencias/Publicaciones/Imagenes/Portadas_Libros/Quimica/Introduccion_al_analisis_estructural_de_proteinas_y_glicoproteinas/Analisis_estructural_proteinas_y_glicoproteinas.pdf)
- Via, A., & Helmer-Citterich, M. (2004). A structural study for the optimisation of functional motifs encoded in protein sequences. *BMC Bioinformatics*, (5), 50. DOI: 10.1186/1471-2105-5-50.
- Wegener, M., & Dietz, K. (2022). The mutual interaction of glycolytic enzymes and RNA in post-transcriptional regulation. *RNA*, 28 (11), 1446 - 1468. DOI: 10.1261/rna.079210.122.
- Wetlaufer, D. (1973). Nucleation, rapid folding, and globular intrachain regions in proteins. *Proceedings of the National Academy of Sciences of the United States of America.*, 70 (3), 697 - 701. DOI: 10.1073/pnas.70.3.697.
- Williams, S., Yin, L., Foley, G., Casey, L., Outram, M., Ericsson, D., Lu, J., Boden, M., & Kobe, B. (2016). Structure and Function of the TIR Domain from the Grape NLR Protein RPV1. *Frontiers in Plant Science*, 7, 1850. <https://doi.org/10.3389/fpls.2016.01850>
- Yang, K., Fang, X., Xu, Y., & Liu, B. (2019). Protein fold recognition based on multi-view modeling, *Bioinformatics*, 35 (17), 2982 – 2990. <https://doi.org/10.1093/bioinformatics/btz040>
- Yu, L., Tanwar, D., Penha, E., Wolf, Y., Koonin, E., & Basu, M. (2019). Grammar of protein domain architectures. *Proceedings of the National Academy of Sciences*, 116 (9), 3636 – 3645. <https://doi.org/10.1073/pnas.1814684116>