

¿Es posible identificar nuevos fármacos desde tu computadora?

Is it possible to identify new drugs from your computer?

María Teresa Alvarado Parra, María Gabriela Mancilla Montelongo y
Karina Verdel Aranda

Resumen

¿Cómo se obtienen los nuevos fármacos? ¿De dónde salen nuevos productos para descontaminar un lago o el suelo? Las plantas y los microorganismos producen millones de moléculas que pueden tener una de estas importantes aplicaciones biotecnológicas. Hasta hace unos años, la caracterización biotecnológica era un proceso costoso y complicado. Sin embargo, la revolución de la bioinformática trajo consigo las bases de datos biológicas que, en conjunto con el desarrollo de la tecnología, han permitido el diseño y desarrollo de algoritmos para conocer la función biológica de un compuesto. Gracias a ello, actualmente se puede predecir la actividad biológica de una gran variedad de moléculas antes de emprender un experimento en el laboratorio. El objetivo de este artículo es compartir cómo es que se pueden caracterizar productos naturales, como fármacos, desde una computadora personal, utilizando bases de datos pertinentes y métodos computacionales *in silico*, para después comprobar la actividad biológica con pruebas *in vitro*.

Palabras clave: base de datos, *in silico*, acoplamiento molecular, productos naturales, bioinformática.

CÓMO CITAR ESTA COLABORACIÓN

Alvarado Parra, María Teresa, Mancilla Montelongo, María Gabriela, y Verdel Aranda, Karina. (2024, enero-febrero). ¿Es posible identificar nuevos fármacos desde tu computadora? *Revista Digital Universitaria (RDU)*, 25(1). <http://doi.org/10.22201/cuaieed.16076079e.2024.25.1.6>

Universidad Nacional Autónoma de México, Coordinación de Universidad Abierta, Innovación Educativa y Educación a Distancia (CUAIEED)

Este es un artículo de acceso abierto bajo la licencia de Creative Commons 4.0



María Teresa Alvarado Parra

Universidad de California, San Diego, USA

Licenciada en Biología de la Universidad Autónoma de Yucatán. Actualmente se dedica a la generación de modelos metabólicos a escala genómica en plantas en la Universidad de California, San Diego, USA. En 2024 recibió la presea Fulbright-García Robles para iniciar sus estudios de posgrado en Estados Unidos. Cuenta con dos años de experiencia en el área de bioinformática y es amante de la programación.

 [0000-0002-7802-4040](https://orcid.org/0000-0002-7802-4040)

 [linkedin.com/in/tere-parra-303ab6166/](https://www.linkedin.com/in/tere-parra-303ab6166/)

 github.com/Tere-Parra

María Gabriela Mancilla Montelongo

Universidad Autónoma de Yucatán (UADY)

Licenciada en Química (UASLP) y Doctora en Ciencias (Química Orgánica, Universidad de Cádiz, España). Es Investigadora CONACYT en la Universidad Autónoma de Yucatán (UADY), su línea de estudio se enfoca en los mecanismos de acción y el valor nutracéutico de productos naturales para el control de parásitos de cabras y borregos. Es miembro del Sistema Nacional de Investigadoras e Investigadores (SNI, nivel 1). Además de las publicaciones científicas y de divulgación, ha dirigido tesis de todos los grados académicos y participa activamente en diversos programas de introducción de jóvenes mujeres a la ciencia.

 [0000-0001-6481-6496](https://orcid.org/0000-0001-6481-6496)

 [researchgate.net/profile/Gabriela-Mancilla-Montelongo](https://www.researchgate.net/profile/Gabriela-Mancilla-Montelongo)

 [linkedin.com/in/gabrielamancillam/](https://www.linkedin.com/in/gabrielamancillam/)

Karina Verdel Aranda

Tecnológico Nacional de México, campus Chiná

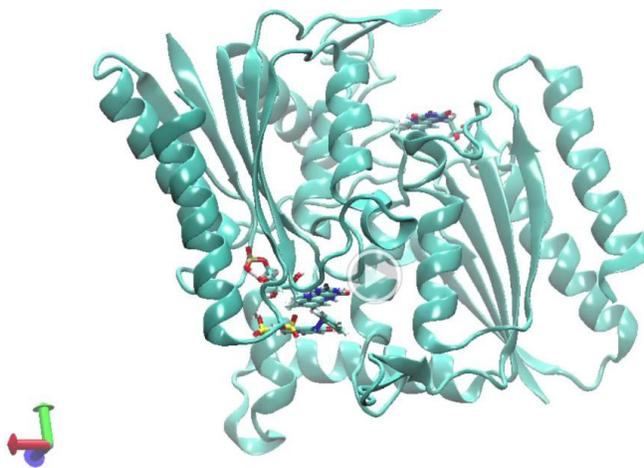
Se dedica a la investigación en el campo de la biotecnología microbiana para resolver problemas agroalimentarios. También siente un gran compromiso con la docencia y la comunicación de los resultados de la ciencia a todo el público. Su investigación está centrada en la caracterización y bioprospección, mediante metagenómica y minería genómica, de comunidades bacterianas en nichos ecológicos del Trópico de México, así como la caracterización genética y bioquímica de clústeres de genes biosintéticos y productos naturales encontrados por minería de genomas de actinobacterias. Estudió Ing. Bioquímica Industrial (UAM-I), Maestría en Ciencias Bioquímicas (IBT-UNAM) y Doctorado en Ciencias en la Especialidad de Biotecnología de Plantas (UGA-CINVESTAV Irapuato). Hizo una Estancia Postdoctoral en la UGA-CINVESTAV Irapuato. Actualmente es docente e investigadora del Tecnológico Nacional de México, campus Chiná.

 [0000-0002-7827-5464](https://orcid.org/0000-0002-7827-5464)

Introducción

Hace un año, trabajé en un proyecto para la biorremediación¹ del río Atoyac, en Puebla, México. Entre los contaminantes de este río, están los tintes usados en la industria textil, que son altamente tóxicos y poco degradables. Para llevar a cabo dicha biorremediación, los investigadores querían utilizar bacterias que produjeran enzimas capaces de reducir estos compuestos tóxicos. El problema era que no estaban seguros si las enzimas producidas por las bacterias eran capaces de reaccionar de manera específica con esas moléculas. Entonces, mi trabajo fue realizar computacionalmente un acoplamiento molecular de las enzimas bacterianas² propuestas con cada una de las moléculas tóxicas seleccionadas. Así, si la afinidad de unión³ entre la enzima y el compuesto era buena, quería decir que la enzima era capaz de reducir el compuesto (ver video 1). Una vez realizada la propuesta, el grupo de investigación recopiló los análisis y diseñó un sistema bacteriano capaz de degradar esos compuestos (Acevedo et al., 2021).

Video 1. Resultado de una simulación dinámica realizada en GROMACS a 10 ns de *Pseudomonas putida* y el compuesto azoico azul índigo. Crédito: el resultado se elaboró y visualizó con el software Visual Molecular Dynamics (vmd).



El anterior testimonio, perteneciente a una estudiante de la Licenciatura en Biología, es la punta del iceberg sobre lo que se puede hacer o predecir desde una computadora. En los últimos años, la demanda por los métodos computacionales en numerosas investigaciones ha ido creciendo y, con ello, la importancia de hacer más accesible la bioinformática y la biología computacional en el descubrimiento de nuevos compuestos activos. Por ello, en este artículo se describen métodos computacionales generales para la caracterización biotecnológica. Esto incluye el concepto de *acoplamiento molecular* de enzimas y ligandos, junto con los pasos que se siguen para llevar a cabo este análisis computacional.

¹ *Biorremediación.* Proceso en el que se utiliza un producto natural (planta, hongo, bacteria, enzima) para eliminar, absorber o transformar un contaminante del medio ambiente.

² *Enzimas bacterianas.* Proteínas producidas por bacterias que aceleran una reacción específica, como la oxidación-reducción.

³ *Afinidad de unión.* Fuerza con la que interactúa una molécula, como una proteína o enzima, con un compuesto en particular, como un contaminante tóxico.

Fuentes de nuevos compuestos y cómo analizar grandes volúmenes de datos

Los productos naturales (PN), en un sentido amplio, se refieren a cualquier parte, biomolécula (como las enzimas o proteínas) o compuesto químico (también llamado metabolito secundario) obtenidos de un ser vivo, como una planta o un microorganismo (Sarker y Nahar, 2012). Una de las aplicaciones más importantes de los PN es en biotecnología ambiental, específicamente en la biorremediación. Las plantas y microorganismos tienen un papel fundamental en el reciclaje de carbono, nitrógeno, fósforo y azufre. Al mismo tiempo, sus PN son capaces de degradar compuestos contaminantes más complejos como los provenientes de las actividades industriales. Las principales aplicaciones de estos PN son el tratamiento de aguas residuales, degradación de xenobióticos⁴ y petroquímicos (Glazer y Nikaido, 2007). Este tipo de moléculas pequeñas, como los contaminantes o los metabolitos secundarios, se convierten en ligandos cuando interactúan con las biomoléculas, como las proteínas.

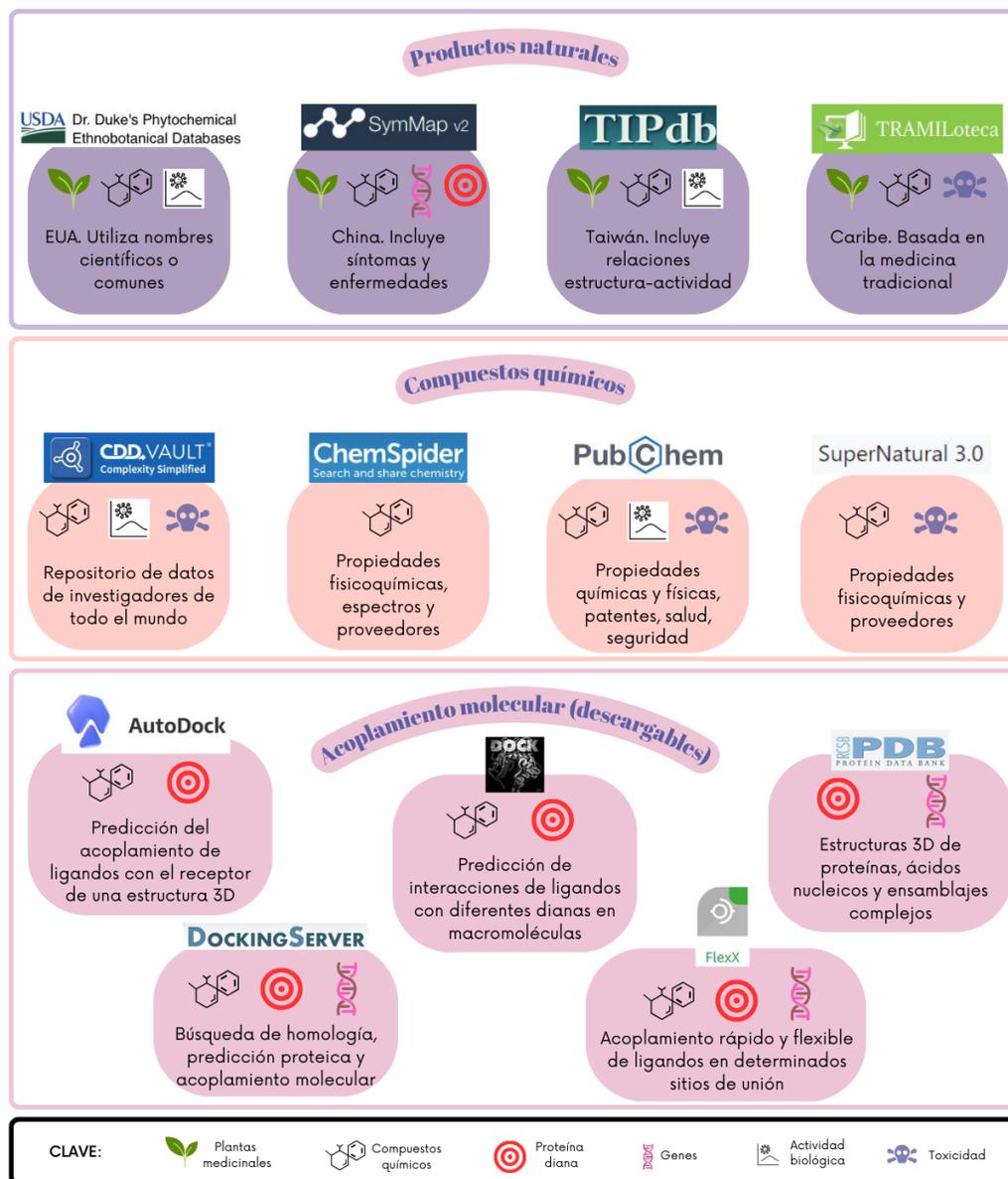
Al recopilar una vasta información taxonómica, molecular, genética, metabólica y toxicológica de PN se ha generado la necesidad de almacenar y poner a disposición grandes volúmenes de datos. El surgimiento de las bases de datos (BD) permitió resolver este problema de almacenamiento y mejoró la disponibilidad de datos biológicos. Hoy en día, la bioinformática es esencial para manejar y comprender esta información. El uso de diferentes métodos que combinan algoritmos, lenguajes de programación, programas y otras herramientas han dado lugar a lo que se conoce como análisis *in silico* (Cañedo-Andalia y Arencibia Jorge, 2004). Esto quiere decir que se analizan en la computadora las interacciones entre diferentes PN. Con la información que nos proveen las BD es posible predecir la actividad de un PN en tres etapas básicas (ver figura 1).

Figura 1. Cómo caracterizar productos naturales activos desde tu computadora, en tres sencillos pasos: 1) conocer las bases de datos, 2) aprender y analizar mediante un método computacional, y 3) corroborar el resultado *in silico* con un experimento *in vitro*. Crédito: imágenes y vectores de Freepik.com.



⁴Xenobióticos. Cualquier sustancia que se encuentra dentro de un organismo pero que no le pertenece o no ha sido producida por él, como fármacos o contaminantes.

Tabla 1. Algunas bases de datos de productos naturales, actividad biológica y acoplamiento molecular.



Paso 1. Bases de datos de productos naturales y de actividad biológica

Las BD de productos naturales cambian de acuerdo con la zona geográfica, tipo de organismo (hierbas, plantas vasculares, cactáceas, etcétera), propiedad biológica, uso y aplicación (antibióticos, anticancerígenos, suplementos alimenticios, entre otros), compuestos químicos o proteínas. Otro tipo de BD son como bibliotecas virtuales que resguardan artículos especializados y permiten acceder a

información relevante para elegir un modelo de estudio. Las BD enfocadas en actividad biológica almacenan propiedades de interacción entre moléculas (compuestos químicos-proteínas-genes). La *actividad biológica* es el potencial que tiene una molécula para llevar a cabo una función, por ejemplo, su uso contra una determinada enfermedad (Jackson et al., 2007). Algunas BD reportan datos como concentraciones inhibitorias, dosis efectivas y dosis letales o tóxicas. La mayoría de las BD son de acceso libre (algunos ejemplos se encuentran en el cuadro 1), otras son exclusivas para investigadores y requieren de un registro. También existen BD con un costo, las cuales, en su mayoría son generadas para la industria farmacéutica.

Paso 2. Métodos computacionales: el acoplamiento molecular

El *acoplamiento molecular* es una predicción de las interacciones entre dos PN: un ligando (molécula pequeña) y una proteína. Los primeros acoplamientos moleculares se basaron en métodos QSAR (relaciones cuantitativas estructura-actividad, por sus siglas en inglés) (Ekins et al., 2007). Estos métodos utilizaban computadoras con procesadores hechos de silicio para estimar la actividad biológica de las moléculas; por ello a estos análisis se les llamó *in silico*. Las primeras predicciones con computadoras usaban modelos matemáticos que relacionaban una estructura con una propiedad fisicoquímica o una actividad biológica, siguiendo como principio la estadística. El primer registro de un fármaco con un diseño *in silico* es el mesilato de nelfinavir, usado para inhibir una proteína del Virus de Inmunodeficiencia Humana (VIH) (Meza Menchaca et al., 2020).

Después surgieron metodologías de acoplamiento basadas en descriptores, reglas, conocimiento, ligandos, objetivos, afinidad virtual, entre muchas otras (Ekins et al., 2007). Dentro de las más usadas se encuentran aquellas que dependen de la disponibilidad de información estructural. Por esta razón, durante un tiempo, los científicos se concentraron en experimentos con las estructuras 3D de proteínas. Todas estas metodologías condujeron a la creación de múltiples interfaces y softwares de afinidad virtual (ver video 1).

Gracias a las numerosas metodologías, se han descubierto inhibidores y dianas terapéuticas que han servido para probar diferentes ligandos que prometen ser futuros fármacos. Es fascinante el control en tiempo récord de brotes de enfermedades a partir de la identificación de compuestos bioactivos y su afinidad a las dianas terapéuticas (Meza Menchaca et al., 2020). Un ejemplo bastante familiar fue la búsqueda de fármacos para el tratamiento de COVID-19, donde el acoplamiento molecular fue “el rey de los métodos computacionales” (Pavan y Moro, 2023).

El acoplamiento molecular ha evolucionado de manera que nos ha permitido reportar y comprobar interacciones de distinta naturaleza, como interacciones ligando-proteína (ver figura 2 y video 1), proteína-proteína, ADN-ligando, ADN-proteína, sólido-proteína, sólido-ligando, entre otras (Dar y Mir, 2017).

Figura 2. Acoplamiento molecular proteína-ligando. La estructura terciaria de la proteína está marcada en verde y la molécula pequeña (ligando) en rojo. Visualización con el programa PyMol donde, utilizando los botones de los apartados marcados, se puede eliminar las moléculas de agua (paso (d)) y observar las interacciones del acoplamiento (paso (g)).



¿Cómo se puede llevar a cabo un análisis de acoplamiento molecular?

La metodología para realizar los acoplamientos moleculares puede resumirse en siete pasos (ver listado a – g), lo que requiere diferentes BD y programas computacionales (ver figura 1). Cabe destacar que esta secuencia puede considerarse como una receta básica general, pero cada bioinformático la modifica y adapta según sus objetivos y lo que intenta analizar.

- a) *Obtención de estructuras.* Consulta de diferentes BD para descargar las estructuras 3D de las moléculas para el acoplamiento (ver figura 1 y tabla 1).
- b) *Búsqueda de dominios funcionales.* Una proteína puede tener dominios funcionales⁵ compartidos con varias familias proteicas, pero esto no significa que mantengan su función biológica. De esta manera, se consultan los dominios funcionales, se anotan y, después del acoplamiento molecular, se verifica si la unión sucedió dentro de estos dominios.
- c) *Exploración de sitios activos.* Las proteínas contienen sitios específicos donde otras moléculas pueden unirse y llevar a cabo su función biológica. A estos sitios se les conoce como *bolsillos de unión* o *sitios activos*. Éstos se muestran como cavidades en las estructuras 3D de las proteínas donde entrará el ligando.
- d) *Minimización energética de moléculas.* A veces las moléculas 3D contienen moléculas de agua u otras con metales pesados. Éstas pueden interferir en el acoplamiento molecular y deben eliminarse. Después, se minimiza la energía de las moléculas 3D para corregir distorsiones o tensiones, reducir el movimiento y garantizar una estructura estable (ver figura 2).

⁵ *Dominios funcionales.*
Secuencias de aminoácidos iguales que pueden encontrarse en proteínas similares a las de nuestra consulta.

- e) *Acoplamiento molecular*. En este paso se seleccionan los aminoácidos que conforman el sitio activo de la proteína y se realiza el acoplamiento molecular de los ligandos (ver figura 2).
- f) *Análisis de resultados*. El acoplamiento molecular arrojará varios resultados, pero los más importantes son la energía de enlace (describe la fuerza de interacción proteína-ligando) y el *score* o puntuación de acoplamiento (refleja la calidad de la unión). Mientras menores sean estos valores, más probable es que la función biológica se cumpla.
- g) *Visualización de interacciones*. Se observan las interacciones intermoleculares (fuerzas de van der Waals, interacciones iónicas, etcétera). También se verifica si la unión sucedió dentro de los aminoácidos que conforman los dominios funcionales del paso b (ver figura 2).

¿Por qué se realizan los pasos de minimización y medición de energías en las biomoléculas? Como puede observarse en el video 1, estas biomoléculas no son rígidas y estáticas como rocas. Su interacción con los ligandos para predecir una posible actividad biológica tampoco es como armar legos. Más bien, tienen unas partes que les dan estructura y otras que son como resortes en movimiento. Entonces, se usan ecuaciones matemáticas para predecir cómo se mueven estos resortes, qué carga tienen y cómo interactúan con los ligandos o las moléculas de agua. Justo con las herramientas computacionales se pueden hacer estos cálculos matemáticos y probar muchas condiciones y ligandos de manera eficiente y muy cercana a la realidad.

Paso 3. ¿Análisis *in silico* o análisis *in vitro*?

Un experimento *in vitro* es aquel donde se comprueba si un PN es activo. Se utilizan placas o cajas plásticas o de vidrio (de ahí el término *in vitro*) y, para que sea válido, se requieren múltiples repeticiones de una misma prueba. Entonces, si ya se realizó el análisis *in silico*, ¿es necesario realizar un experimento *in vitro*? ¿Cuándo es necesario hacer un análisis *in silico* o uno *in vitro*? Es normal plantearse estas preguntas y la respuesta es más sencilla de lo que parece. Cuando se reporta un nuevo compuesto con determinada actividad para resolver un problema médico o ambiental es común ver un análisis *in vitro*, en el que el compuesto fue probado a diferentes concentraciones. Sin embargo, los análisis *in vitro* no revelan la interacción que existe entre los compuestos y, por tanto, no tenemos conocimiento sobre por qué es efectivo y las posibles alternativas. Las investigaciones que usan herramientas *in silico* permiten entender, en un menor tiempo, la interacción entre compuestos. Por lo tanto, el análisis *in silico* es un complemento de los análisis *in vitro*.

El futuro está en las computadoras

En la actualidad, se puede utilizar una computadora personal para dar los primeros pasos en el descubrimiento de compuestos activos a través del acoplamiento molecular. Además, gracias a internet, podemos acceder vía remota a múltiples BD y programas bioinformáticos. Sin embargo, ésta es una visión reducida del potencial que tienen los métodos computacionales. En análisis masivos de datos *in silico* es ideal la independencia de la web. Para esto se usan computadoras de alta gama que tienen miles de GB de memoria, cientos de TB de almacenamiento y decenas de núcleos de procesamiento. Estos equipos suelen estar conectados entre sí formando *clusters*⁶, que pueden analizar cientos de procesos bioinformáticos al mismo tiempo de manera eficiente.

Conclusión

Aún nos queda mucho por descubrir, pero, gracias a la bioinformática y la biología computacional, los descubrimientos están siendo mucho más rápidos que nunca. En general, los bioinformáticos descubren fármacos desde su computadora en tres pasos: a) conociendo BD sobre organismos y compuestos activos, b) usando métodos computacionales como el acoplamiento molecular y c) combinando métodos *in silico* e *in vitro*.

¿Cuáles son las nuevas tendencias y el futuro de esta área de la biología computacional? Debido a la gran cantidad de datos que actualmente existen, conocidos como *big data*, los investigadores y las compañías farmacéuticas ya están usando inteligencia artificial (IA) para analizar toda esta información molecular. Lo anterior les permite predecir de manera muy precisa nuevas estructuras, lo que ahorra tiempo y dinero. Además, el avance en el desarrollo de métodos computacionales en general se enfoca en abordar los principales retos en esta área, como la alta flexibilidad de las biomoléculas y la lentitud en la unión-disociación de las mismas. Sin embargo, hay temas que aún necesitan respuesta, como el entendimiento de los mecanismos de reacción, los procesos termodinámicos y la cinética de las interacciones moleculares, entre otros.

Referencias

- ❖ Acevedo, A., Alvarado, M., Castañón, V., y Diaz, V. (2021). *Azoyac: Cleaning Atoyac's water using synthetic biology*. CONCYTEP. <https://app.jogl.io/es/project/762/Azoyac>
- ❖ Cañedo Andalia, R., y Arencibia Jorge, R. (2004). Bioinformática: en busca de los secretos moleculares de la vida. *Acimed*, 12(6), 1-29. <http://tinyurl.com/bde2upvu>
- ❖ Dar, A. M., y Mir, S. (2017). Molecular Docking: approaches, types, applications and basic challenges. *Journal of Analytical & Bioanalytical Techniques*, 8, 356. <https://doi.org/10.4172/2155-9872.1000356>

⁶ *Clusters*. Grupo de computadoras que están conectadas y trabajan en conjunto como si fueran una sola.

- ❖ Ekins, S., Mestres, J., y Testa, B. (2007). *In silico* pharmacology for drug discovery: methods for virtual ligand screening and profiling. *British Journal of Pharmacology*, 152(1), 9–20. <https://doi.org/10.1038/sj.bjp.0707305>
- ❖ Glazer, A. N., y Nikaido, H. (2007). *Microbial biotechnology: fundamentals of applied microbiology* (2.ª ed.) Cambridge University Press.
- ❖ Jackson, C. M., Esnouf, M. P., Winzor, D. J., y Duewer, D. L. (2007). Defining and measuring biological activity: applying the principles of metrology. *Accreditation and Quality Assurance*, 12(6), 283-294. <https://doi.org/10.1007/s00769-006-0254-1>
- ❖ Meza Menchaca, T., Juárez-Portilla, C., y Zepeda, R. C. (2020). Past, present, and future of molecular docking. En V. Gaitonde, P. Karmakar, y A. Trivedi (Eds.), *Drug discovery and development – new advances* (cap. 2). IntechOpen. <https://doi.org/10.5772/intechopen.90921>
- ❖ Pavan, M., y Moro, S. (2023). Lessons Learnt from COVID-19: computational strategies for facing present and future pandemics. *International Journal of Molecular Sciences*, 24(5), 4401. <https://doi.org/10.3390/ijms24054401>
- ❖ Sarker, S. D., y Nahar, L. (2012). *Natural products isolation. Methods in Molecular Biology* (3.ª ed., vol. 864). Humana Press. <https://doi.org/10.1007/978-1-61779-624-1>