



EDITORIAL
*Premio Nobel de Química 2024 en
Diseño y Estructura de la Proteínas*

EDITORIAL

SOBRE EL PREMIO NOBEL DE QUÍMICA 2024 EN DISEÑO Y ESTRUCTURA DE LAS PROTEÍNAS

El otorgamiento del Premio Nobel de Química 2024 a David Baker, Demis Hassabis, y John M. Jumper es un reconocimiento al campo de la biología computacional, en particular, en la predicción computacional de estructuras proteicas. Este reconocimiento destaca el impacto revolucionario de las tecnologías de aprendizaje automático e inteligencia artificial en la solución de problemas biológicos complejos que, en tiempos pasados, parecían intratables. A través de este premio, se enfatiza cómo la confluencia de la biología con tecnologías avanzadas puede despejar misterios del mundo natural que han desafiado a los científicos durante décadas.

Las proteínas son moléculas fundamentales en todos los procesos celulares y desempeñan sus funciones críticas a través de su estructura tridimensional. El “problema del plegamiento de proteínas”, es decir, la predicción de estructuras proteicas a partir de secuencias de aminoácidos, ha constituido uno de los desafíos más grandes de la biología molecular.

A partir de la propuesta de los patrones de plegamiento de las alfa hélices, por Linus Pauling, Robert Corey, y Herman Branson en 1951 y de las hojas beta plegadas por Pauling y Corey, también en 1951, se fue encontrando que estos patrones se repetían en la inmensa mayoría de las proteínas y que además había otros tipos de plegamientos, asas y vueltas que tendían a conectar esas regiones. Esos patrones indicaban que la estructura de las proteínas no era al azar, pero no daban ninguna pista de cómo poder predecir en forma detallada, a partir de la secuencia de aminoácidos, cual estructura finalmente adoptaría la proteína.

Tradicionalmente, resolver las estructuras proteicas dependía de técnicas experimentales como la cristalografía de rayos X, la resonancia magnética nuclear, y, más recientemente, la criomicroscopía

electrónica que son muy efectivas, pero son laboriosas y requieren tiempo para realizarse.

A lo largo de muchos años esto permitió la creación de bases de datos muy confiables y sólidas con estructuras de proteínas determinadas experimentalmente como el *Protein Data Bank* (PDB), que contiene un número muy grande de estructuras tridimensionales asociadas a secuencias de aminoácidos y que pueden ser utilizadas por computadoras y programas para el procesamiento de datos.

El desarrollo de métodos computacionales cada vez más complejos y sofisticados ha traído una revolución a este campo, permitiendo predicciones más rápidas y precisas, lo que ha cambiado las reglas y los procedimientos para la investigación y las aplicaciones prácticas.

Una manera en que los científicos habían estado intentando resolver “el problema del plegamiento de las proteínas”, fue a partir de principios físicos y químicos. Para hacerlo de esta forma, hay que calcular computacionalmente valores de parámetros como ángulos torsionales, puentes de hidrógeno, interacciones electrónicas pi-pi, energías de tensión de enlaces y otras variables, tratando de alcanzar los mínimos de energía. Estos cálculos son una tarea enorme, y más, si se trata de una proteína grande, por lo que siempre se buscaban atajos para simplificar y acortar los cálculos.

Las innovaciones aportadas por Baker, Hassabis, y Jumper han sido muy importantes en este cambio para abordar el problema. El desarrollo, por parte de estos científicos, de algoritmos avanzados que aprovechan enormes cantidades de datos biológicos, como los contenidos en el PDB, y recursos computacionales potentes, ha dado lugar a avances en la predicción de estructuras proteicas con mucha precisión. Utilizan avances en algoritmos de aprendiza-

dizaje de máquina aplicados a secuencias de proteínas y patrones de estructuras de proteínas con un poder y una profundidad que no se había alcanzado antes.

AlphaFold 2 de Google DeepMind, liderado por Hassabis y Jumper, y RoseTTAFold del equipo de Baker en la Universidad de Washington, son dos ejemplos destacados de este tipo de esfuerzos computacionales. Utilizando técnicas de aprendizaje profundo, estas herramientas interpretan el lenguaje complejo de las secuencias proteicas y predicen sus estructuras tridimensionales con una precisión asombrosa, evitando los cálculos químicos y físicos del plegamiento mencionados antes. Estos programas aprenden de una base de datos extensa de estructuras proteicas conocidas y detectan patrones y relaciones que predicen el resultado final del plegamiento. Esta capacidad para predecir la configuración física de proteínas a partir de sus secuencias solas ha transformado fundamentalmente el campo de la biología estructural.

La exactitud de AlphaFold 2 se validó en la competencia bianual de Evaluación Crítica de la Predicción de Estructuras Proteicas (CASP), la cual proporciona una plataforma rigurosa para medir el progreso en la predicción de estructuras proteicas. La eficacia de AlphaFold 2 en la competencia CASP14 en noviembre de 2020 demostró su gran supremacía, a menudo igualando o superando los métodos tradicionales en términos de velocidad y precisión. Esto hizo que el jurado de la competencia declarara que el problema de la predicción de la estructura de las proteínas a partir de su secuencia primaria de aminoácidos estaba básicamente resuelto.

Unos 20 años antes del trabajo de DeepMind sobre AlphaFold2, David Baker y sus colaboradores desarrollaron una herramienta de software que llamaron Rosetta, la cual modelaba estructuras de proteínas usando principios físicos. Inicialmente, Baker utilizó Rosetta para predecir estructuras de proteínas y compitió en varios CASP a lo largo de varios años. Sin embargo, cuando Baker notó la dominancia en la predicción de proteínas de AlphaFold, decidió invertir la función de Rosetta y utilizarla para diseñar proteínas completamente nuevas; algunas de ellas jamás habían sido producidas por la naturaleza. El trabajo de Baker y colaboradores continúa hasta hoy en día, y el comité Nobel lo premió por lograr construir “proteínas espectaculares” haciendo “creaciones” que pueden ser utilizadas como “productos farmacéuticos, vacunas, nanomateriales y pequeños sensores”.

Además de su impacto en la ciencia básica, las implicaciones de estas herramientas computacio-

nales son vastas y se extienden a aplicaciones prácticas en numerosos campos. Uno de los que más se mencionan es el descubrimiento de fármacos. Y, aunque la capacidad de predecir rápidamente estructuras proteicas es fundamental para el diseño de medicamentos efectivos, el tener esas estructuras no es equivalente a tener una molécula que afecte su función. Ni tampoco siempre acelerará el proceso de diseño de fármacos, como muchas personas creen y/o esperan que lo haga.

Desde hace bastantes años ha existido la creencia de que nuestra deficiencia en conocimiento sobre la estructura de las proteínas nos está impidiendo dar grandes pasos en el descubrimiento de fármacos. Pero esta creencia revela que no se tienen en cuenta otros factores que también son muy importantes en el proceso, después de haber resuelto el primer paso, que es conocer su estructura. Las proteínas tienen movimientos y cambian de forma conforme llevan a cabo sus funciones y, también, a veces, cuando se unen a moléculas farmacológicas. La energética de esos eventos depende de múltiples factores, sobre los cuales muy pequeñas diferencias pueden tener un efecto muy significativo. Además, el diseño de fármacos, que modulen efectivamente la función de las proteínas, requiere una comprensión profunda de las interacciones moleculares y las fuerzas que las regulan. Por lo tanto, aunque las herramientas computacionales han mejorado nuestras capacidades de manera espectacular de trabajar en esto, son solo una parte de un conjunto más amplio de técnicas que deben integrarse para abordar los desafíos complejos del desarrollo de fármacos. Como ya se dijo, el conocimiento de las estructuras de las proteínas es un muy buen primer paso para lograr este fin.

La contribución de los ganadores del Premio Nobel de Química 2024 va más allá de resolver un problema biológico involucrando muchos datos. Han abierto nuevas avenidas para la investigación y la innovación, mostrando el poder de los enfoques interdisciplinarios que combinan biología, ciencia computacional y física para investigar el mundo natural. A medida que estas herramientas computacionales continúen evolucionando y mejorando, su potencial para impactar la ciencia y la medicina se expandirá aún más, prometiendo descubrimientos nuevos e innovaciones que, por ahora, siguen fuera de nuestro alcance.



Ruy Pérez Montfort
Instituto de Fisiología Celular
Universidad Nacional Autónoma de México
ruy@ifc.unam.mx