

El Protein Data Bank (PDB) y su impacto en la investigación científica

Enrique Rudiño-Piñera¹ **iD**, Verónica Quintero-Hernández **iD**, Victor Rivelino Juárez-González* **iD**.

¹Departamento de Medicina Molecular y Bioprocesos, Instituto de Biotecnología, Universidad Nacional Autónoma de México, Av. Universidad 2001, Col. Chamilpa, Cuernavaca, Morelos 62210 México.

²CONACYT-Laboratorio de Ecología Molecular Microbiana, Centro de Investigaciones en Ciencias Microbiológicas, Instituto de Ciencias, Benemérita Universidad Autónoma de Puebla (BUAP), Puebla, México.

*Email autor correspondiente: *victor.juarez@ibt.unam.mx

Recibido: 19 octubre 2021. **Aceptado:** 10 diciembre 2021

RESUMEN

El *Protein Data Bank* (banco de datos de proteínas), abreviado como PDB, es el banco de información estructural de proteínas a nivel atómico, donde están depositadas todas las estructuras tridimensionales de las macromoléculas biológicas determinadas por la comunidad científica mundial y que incluye a proteínas, ácidos nucleicos, carbohidratos y lípidos. Su uso es libre y hasta noviembre del 2021 tiene 184,202 estructuras depositadas. La mayor parte de la información tridimensional con la que cuenta el PDB con respecto a las biomoléculas es la referente a proteínas con 158,644 estructuras depositadas y la técnica más utilizada para la determinación de las estructuras es la difracción de rayos X con 161,144 estructuras depositadas. En este año, el PDB cumple 50 años y la información que nos proporciona ha revolucionado la visión que teníamos sobre los procesos biológicos a nivel atómico. El PDB es una enorme base de datos de estructuras y una fuente de información importante sobre diversas macromoléculas, las cuales pueden ser analizadas por una diversidad de profesionales como: cristalógrafos, químicos, biólogos, bioquímicos, microbiólogos, médicos, farmacéuticos, entre otros. El conocer la estructura de una macromolécula de interés puede ayudar a entender una gran cantidad de procesos químicos, biológicos y enfermedades, ya que entender a detalle atómico la alteración química y estructural de una macromolécula permite descifrar los mecanismos que participan en el proceso químico o biológico de estudio o en una enfermedad y el daño en el organismo; además se puede realizar el diseño de fármacos para su tratamiento.

Palabras clave: PDB; Difracción de rayos X; RMN; Criomicroscopía electrónica.

ABSTRACT

The Protein Data Bank, abbreviated as PDB, is the structural information bank of proteins at the atomic level, where all the three-dimensional structures of biological macromolecules determined by the world scientific community are deposited and which includes proteins, nucleic acids, carbohydrates, and lipids. Its use is free, and until november 2021 it has 184,202 structures deposited. Most of the three-dimensional information that the PDB has with respect to biomolecules refers to proteins with 158,644 structures deposited and the most used technique for the determination of structures is X-ray diffraction with 161,144 structures deposited. In this year, the PDB celebrates 50 years, and the information it provides us has revolutionized the vision we had about biological processes at the atomic level. The Protein Data Bank is a very large database of structures and is also a source of much important information on various macromolecules, which can be analyzed by a variety of professionals such as: crystallographers, chemists, biologists, biochemists, microbiologists, doctors, pharmacists, among others. Knowing the structure of a macromolecule of interest can help to understand a large number of chemical and biological processes and diseases, since understanding in atomic detail the chemical and structural alteration of a macromolecule allow us to decipher mechanisms that participate in the chemical or biological process or the one that cause disease and damage in the organism; In addition, the design of drugs for its treatment can be carried out.

Keywords: PDB; X-ray diffraction; NMR; Electronic cryomicroscopy.

INTRODUCCIÓN

El PDB es el banco de información estructural de proteínas a nivel atómico, donde están depositadas todas las estructuras tridimensionales (3D) de las macromoléculas biológicas determinadas por la comunidad científica mundial (proteínas, ácidos nucleicos, carbohidratos y lípidos) [1,2]. Su uso es libre y las direcciones web de acceso, incluyendo una que da información para comprender la trascendencia del PDB, son: www.pdb.org (wwPDB: Worldwide Protein Data Bank y RCSB PDB: Homepage) y <https://pdb101.rcsb.org/learn/guide-to->

[understanding-pdb-data/introduction](https://www.rcsb.org/stats/growth/growth-released-structures). Esta enorme fuente de información estructural tuvo sus orígenes en el año de 1971, año en el cual contaba con solo siete estructuras depositadas, entre las cuales se encontraba la Mioglobina (código PDB:1MBN) [3]. Desde ese entonces hasta noviembre de 2021 ha tenido un crecimiento enorme ya que ahora cuenta con 184,202 estructuras depositadas (Figura 1, <https://www.rcsb.org/stats/growth/growth-released-structures>). Justo este año, el PDB se viste de manteles largos ya que cumple 50 años, medio siglo, en que ha revolucionado la visión que teníamos sobre los procesos biológicos a nivel atómico (Figura 2). La determinación de

la estructura tridimensional de una macromolécula, a nivel atómico, se genera con poderosas técnicas desarrolladas por muchos científicos como son: la difracción de rayos X sobre cristales de proteína, utilizando la energía de los rayos X de alta brillantez (fotones/segundo/mm²) para difractar cristales proteicos que contienen aproximadamente 1×10^9 - 1×10^{10} moléculas de proteína, con un arreglo periódico y ordenado dentro del cristal [4–6]; la resonancia magnética nuclear o RMN que usa la información procedente de los núcleos de todos los átomos y los somete a campos magnéticos en los que se encuentra la

macromolécula en solución causando cambios en el espín nuclear [7,8]; la criomicroscopía electrónica, que es capaz de generar imágenes o fotografías a detalle atómico de las macromoléculas, gracias a sus poderosos detectores, procesos de superposición de imágenes y la interacción de las muestras con un haz de electrones, mejorando la resolución de las macromoléculas enfriándolas a temperaturas bajo cero [9,10]. Adicionalmente, existen otras aproximaciones que combinan a las tres técnicas descritas y que permiten conocer las estructuras de macromoléculas difíciles de determinar.

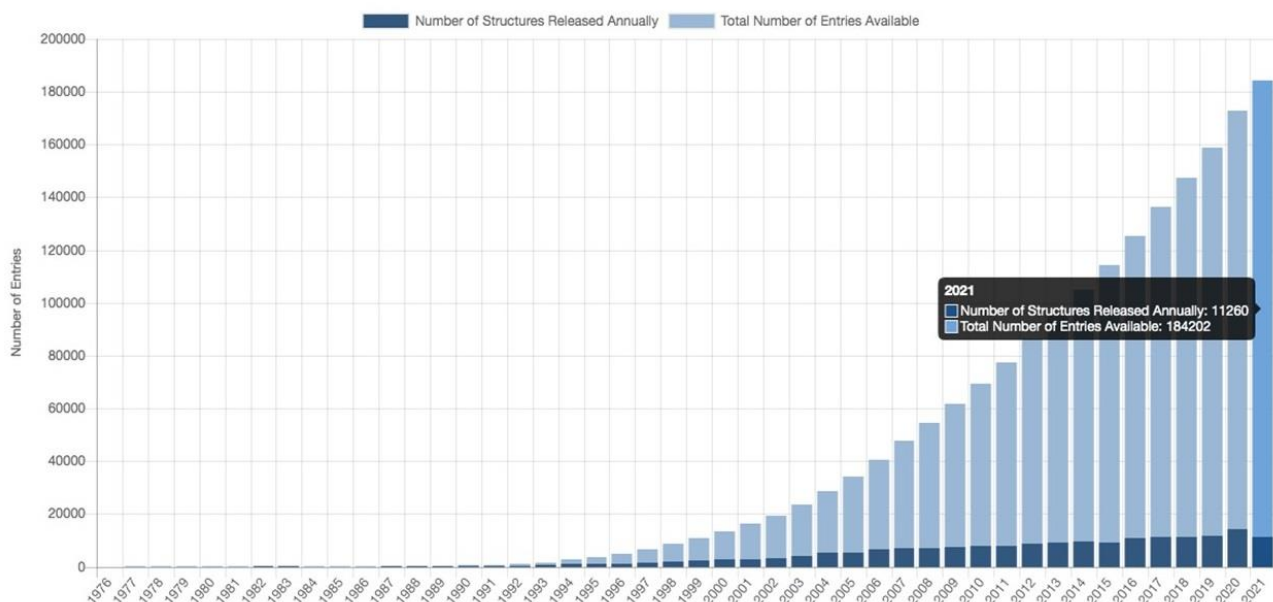


Figura 1. Estadística de las estructuras de macromoléculas depositadas en el PDB por año (1976-2021) determinadas por cualquier técnica. Las barras de color azul claro indican el número total de las estructuras depositadas por año y las barras de azul marino indican los nuevos depósitos de cada año. Imagen tomada de <https://www.rcsb.org/stats/growth/growth-released-structures>.



Figura 2. Imagen conmemorativa de la celebración de los 50 años del PDB o banco de datos de proteínas de 1971-2021. Figura tomada de <https://cdn.rcsb.org/wwpdb/docs/pdb50yrs2021/PDB50-long.png>

Analizando el PDB en este año, se observa que la mayor parte de la información tridimensional con la que contamos con respecto a las biomoléculas, es la referente a proteínas. Salta a la vista que la difracción de rayos X es la técnica más empleada para hacer estas descripciones con 161,144 depósitos (Figura 3, <https://www.rcsb.org/stats/growth/growth-xray>), la segunda técnica utilizada es la RMN con 13,545 depósitos (<https://www.rcsb.org/stats/growth/growth-nmr>), aunque la criomicroscopía electrónica tomará este lugar en los próximos años con 9,210 depósitos (<https://www.rcsb.org/stats/growth/growth-em>); mientras que los métodos múltiples, o métodos combinados, tienen 194 depósitos (<https://www.rcsb.org/stats/growth/growth-multi-method>). Como se mencionó anteriormente, si analizamos el número de estructuras más abundantes en el PDB se observa que corresponde a las proteínas con 160,861 depósitos ([\[protein\]\(https://www.rcsb.org/stats/growth/growth-protein\)\), en segunda posición aparecen los complejos de proteínas con los ácidos nucleicos con 9,823 depósitos \(<https://www.rcsb.org/stats/growth/growth-protein-na-complex>\), en tercera posición aparecen las estructuras del DNA con 2,118 depósitos \(<https://www.rcsb.org/stats/growth/growth-dna>\) y en cuarta posición quedan las estructuras del RNA con 1,571 depósitos \(<https://www.rcsb.org/stats/growth/growth-rna>\).](https://www.rcsb.org/stats/growth/growth-</p></div><div data-bbox=)

Existe una riqueza de información estructural contenida en el PDB, y cualquier persona con acceso a una computadora y a internet tiene acceso libre. Es importante visualizar y analizar las estructuras de estas biomoléculas, pues con ello se puede obtener información muy valiosa que toma mayor relevancia cuando te haces una pregunta biológica, química o médica, entre otras. Considerando que el PDB ésta alimentado en su mayoría por estudios estructurales que provienen, como lo mencionamos arriba, de la difracción de rayos

X, la RMN y la criomicroscopía electrónica, tomaremos 3 ejemplos de este universo de estructuras, para ver más detalles de la información estructural contenida en el PDB.

Determinación de una estructura de proteína mediante difracción de rayos X

El primer ejemplo se centra en el hecho de que durante varios años **diversos** grupos hemos estado trabajando en la resolución de la estructura tridimensional, por medio de la difracción de rayos X, de mutantes generadas en la proteína verde fluorescente o GFP (*green fluorescent protein*) aislada de la medusa *Aequorea victoria* [11], con la idea de explicar estructuralmente porque las mutantes de GFP generadas por evolución dirigida son más fluorescentes, estables o tienen mayor expresión, que la proteína GFP original de la

cual partieron [12,13]. Eventualmente, estas nuevas mutantes serán introducidas en bacterias para ver la localización y migración de otras proteínas de interés científico en las bacterias. De hecho, en el año 2008, se entregó el premio Nobel de química a tres investigadores: Osamu Shimomura [5], Martin Chalfie [14] y Roger Tsien [15], por sus importantes descubrimientos sobre la proteína GFP, la cual tiene la importante propiedad de generar luz verde de manera natural. Esta propiedad permitió el avance de manera exponencial en distintas áreas de investigación como la biología, microscopía, medicina, etc. La proteína GFP, en conjunto con otros marcadores o proteínas de interés, ha permitido conocer, seguir y analizar muchos procesos biológicos [16–18].

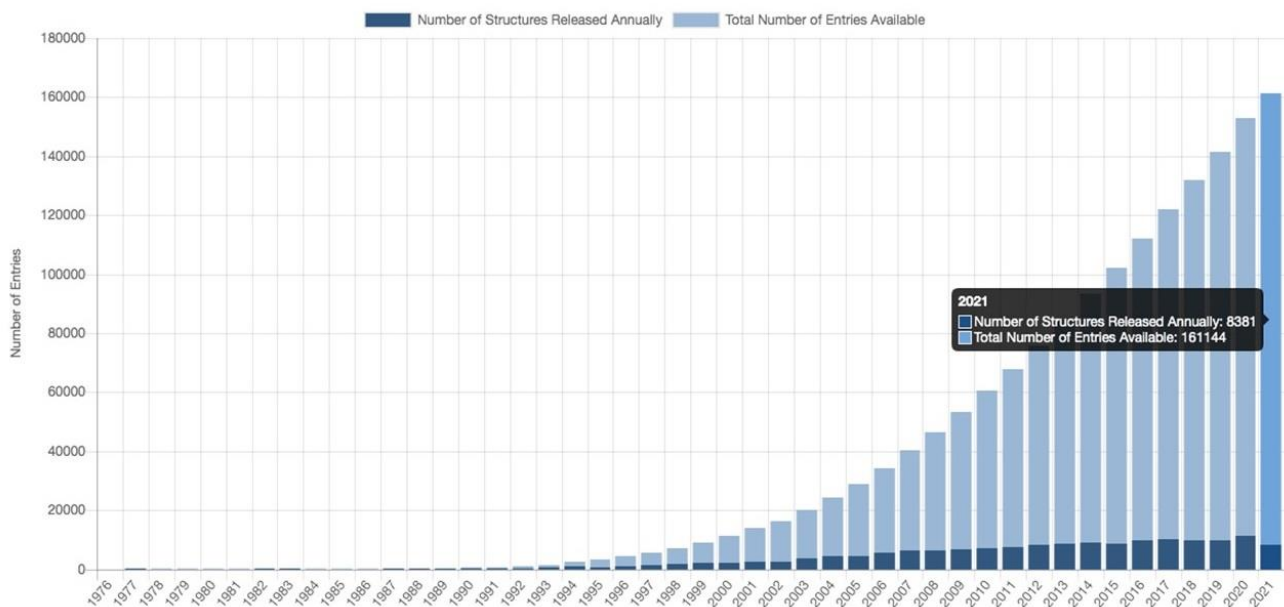


Figura 3. Estadística de las estructuras de todas las macromoléculas depositadas en el PDB por año (1976-2021), por Difracción de rayos X, las barras de color azul claro es el total de estructuras depositadas y las barras de azul marino son los depósitos nuevos de cada año. Imagen tomada de <https://www.rcsb.org/stats/growth/growth-xray>.

En 1996, Roger Tsien, fue capaz de develar la estructura tridimensional de la GFP a una resolución de 1.9 Å (Armstrong), por medio de la cristalografía y difracción de los rayos X, observando a esta resolución un mayor detalle de los átomos que constituyen los aminoácidos de su estructura. Los cristalógrafos son científicos expertos en la resolución de la estructura tridimensional de las macromoléculas, utilizan los rayos X sobre los cristales de proteína y se requiere obtener una alta resolución de la estructura para conocerla a mayor detalle. Haciendo una analogía de esto: es como si observáramos al volcán Popocatepetl desde Puebla, Cuernavaca, o la Ciudad de México; a esa distancia solo se ve la forma piramidal del volcán y si tiene o no nieve, dependiendo de las estaciones del año, o si presenta o no fumarolas, pero no podemos ver más detalle debido a la distancia en que nos encontramos, pero si nos acercamos al volcán y estamos a un lado de él o en las faldas del mismo, podremos ver el tipo, el tamaño y los colores de las rocas y si lo subimos podemos ver el cráter, sus dimensiones, los minerales que lo forman y el lugar de donde salen las fumarolas, por lo tanto, podemos ver más detalle del volcán, e incluso sentir el calor que emite. Siguiendo la analogía, un cristalógrafo solo ve la superficie de la macromolécula cuando la resolución está entre valores de 3 a 10 Å, pero si el cristalógrafo aumenta la resolución en su cristal a valores de entre 2.0 y 1.4 Å puede ver más detalle, incluso detectar el cambio en nubes electrónicas o visualizar átomos de hidrógeno (si bien, en el laboratorio,

este aumento de resolución se debe a que tan exitoso fue el científico para tener cristales cada vez más ordenados). De esta manera Roger Tsien fue capaz de encontrar que la estructura tridimensional de la GFP [19], está formada por un barril de 11 hebras beta antiparalelas conectadas entre ellas, arriba y abajo, por asas o *loops*, en el centro del barril hay una alfa-hélice, y en ella, se encuentra un cromóforo formado por tres aminoácidos S(serina)-Y(tirosina)-G(glicina) en las posiciones 65-67 respectivamente (Figura 4). El cromóforo es la molécula responsable de absorber luz a una longitud de onda determinada y de emitirla en otra, es decir absorbe luz azul y produce luz verde, la característica más importante de estas proteínas y la razón por la que se estudian con tanto detalle (Figura 5).

Si exploramos al banco de datos estructurales para poder revisar la información al respecto de la estructura de la proteína verde fluorescente, necesitamos conocer los 4 caracteres del código de PDB de la estructura de la GFP de la medusa *Aequorea victoria*: 1EMA [5], ese código y la siguiente <https://www.rcsb.org/structure/1EMA> nos dirigen a una ventana, la cual nos arroja información sobre las características estructurales y experimentales de la estructura de la proteína verde fluorescente. Este es un ejemplo de lo que podemos encontrar como usuarios en una estructura determinada por difracción de rayos X, el conocimiento de la estructura tridimensional de la proteína verde fluorescente revolucionó muchas áreas de investigación a nivel mundial, ya que fue posi-

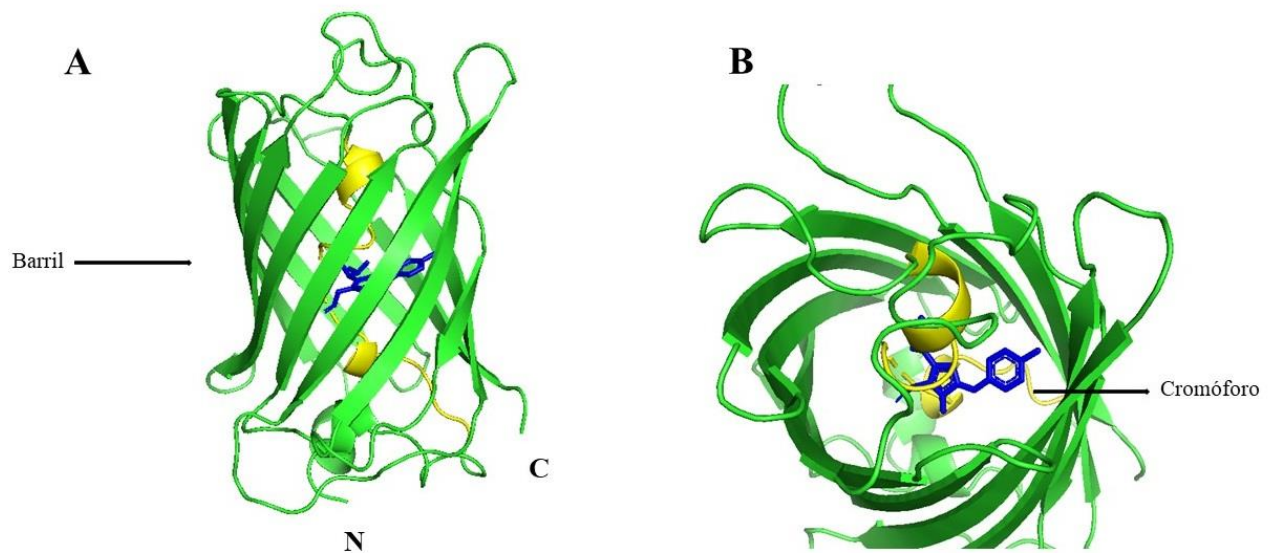


Figura 4. Estructura de la proteína GFP código de PDB: 1EMA. A) Vista lateral y B) vista aérea, se muestran las 11 hebras Beta antiparalelas en color verde, en amarillo está la alfa hélice y el cromóforo está en color azul vistas en PyMOL. Figura realizada con la información de <https://www.rcsb.org/structure/1EMA> y modificada en PyMOL.

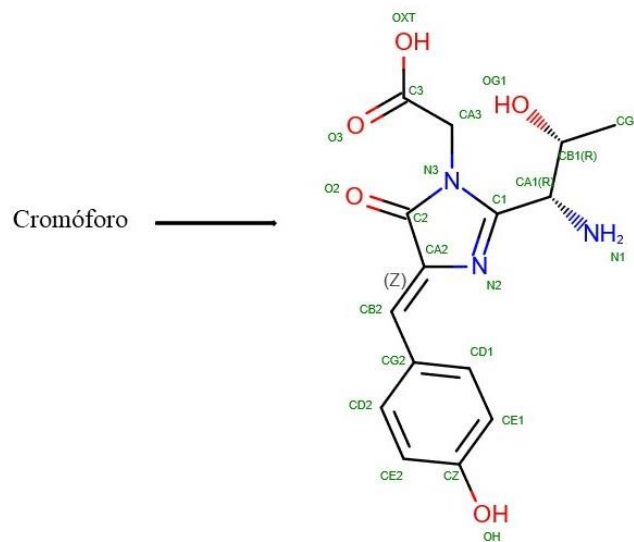


Figura 5. Estructura del cromóforo de la proteína GFP (código de PDB: 1EMA) obtenida del PDB de la sección de moléculas pequeñas. Figura tomada de <https://www.rcsb.org/structure/1ema>.

ble ver muchos procesos biológicos en las células y avanzar en el entendimiento de los mismos.

Determinación de una estructura de proteína mediante resonancia magnética nuclear

En el segundo ejemplo a analizar mencionaremos proteínas que desde hace

mucho tiempo varios grupos mexicanos han estado trabajando con el fin de generar anticuerpos completos tipo IgG (150 kDa) [20,21], Fab (50 kDa) [22] y scFv [21] de última generación o fragmento variable de cadena sencilla (30 kDa) y usarlos con el fin de neutralizar algunos venenos de alacranes en México. Un modelo que destaca es el *Centruroides noxius*, endémico del estado de Nayarit y considerado el alacrán más peligroso para los humanos (Figura 6 A). Al caracterizar su veneno por diferentes tipos de cromatografías (técnicas usadas para la separación de las proteínas y otras moléculas en base a sus propiedades fisicoquímicas, bioquímicas, etc.), se sabe que la toxina más abundante (6.8 %) del veneno total de este alacrán y la responsable de la mayoría de los síntomas del envenenamiento es la toxina Cn2 [20]. Esta toxina presenta una dosis letal 50 (LD₅₀) de 0.25 µg/20 g de peso de ratón hembra CD1 (es decir, que con esa dosis muere la mitad de los ratones inyectados). La estructura tridimensional de Cn2 fue determinada por RMN [23], demostrando la presencia de tres hebras beta antiparalelas y una alfa hélice, estabilizadas por 4 puentes disulfuro (Figura 6 B). Actualmente ya se conocen 7 diferentes anticuerpos que son capaces de neutralizar sus efectos en un 99%, de los cuales los autores de este manuscrito hemos participado en la generación de 4 anticuerpos, uno en formato scFv de 30 kDa [21], 3 en formato Fab de 50 kDa [22], adicionalmente, ya se conocía uno en

formato IgG [20] y hay otros en formatos scFv humanos [24,25]. Con solo neutralizar a la toxina Cn2 es suficiente para contrarrestar el efecto de todo el veneno del alacrán *Centruroides noxius*, por lo cual se le conoce como un veneno simple (compuesto por 70 toxinas diferentes, sales y moléculas pequeñas, etc.). Debido a su efecto mortal, es muy importante neutralizar los venenos de alacranes peligrosos para los humanos y también neutralizar los venenos de otros organismos peligrosos, como pueden ser: serpientes, ranas, caracoles marinos, medusas, etc. Los antivenenos, por lo tanto, disminuyen el número de casos de intoxicaciones, y sobre todo, de defunciones. Si exploramos el código de PDB para la toxina Cn2 el cual es: 1CN2 [23] y el link <https://www.rcsb.org/structure/1CN2>, aparece una ventana que tiene información sobre 15 estructuras tridimensionales, características de su determinación por RMN, el peso molecular de la toxina y mucha más información sobre la caracterización de estas estructuras. Este es un ejemplo de lo que podemos encontrar como usuarios en una estructura determinada por RMN. Por lo anterior, la determinación de la estructura tridimensional de la toxina Cn2 del alacrán *C. noxius* ha proporcionado información muy valiosa, como la identificación de la región de la toxina que es reconocida por los anticuerpos neutralizantes y esto permitirá obtener antivenenos más eficientes.

A)



B)

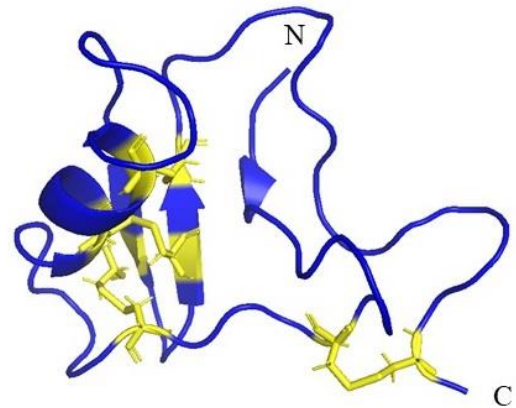


Figura 6. A) El *Centruroides noxius* es el alacrán más peligroso en México, es de color negro rojizo, habita en Narayít. B) Estructura de la toxina Cn2 determinada por NMR (código de PDB:1CN2), en color azul se observan las tres hebras beta antiparalelas, una alfa hélice y los 4 puentes disulfuro, en color amarillo vistas en PyMOL. Imagen del alacrán tomada de <https://enciclovida.mx/especies/88164-centruroides-noxius>. Figura de la toxina realizada con la información de <https://www.rcsb.org/structure/1CN2> y modificada en PyMOL.

Determinación de una estructura de proteína mediante criomicroscopía electrónica

Como tercer ejemplo del universo de estructuras depositadas en el PDB, y recordando la pandemia en que nos encontramos, es muy importante conocer cómo podemos generar vacunas y con ellas anticuerpos que nos protejan de la infectividad del virus que ocasiona la enfermedad COVID-19. El número de muertes provocadas por el SARS-CoV-2 el mundo es de 5.11 millones y en México es de 291,000, desde el primer reporte en Wuhan, China, hasta hoy (hasta el 17 de noviembre del 2021) [26]. Se sabe que las personas que tienen el virus lo pueden propagar con solo estornudar (el virus se transporta en las microgotas de saliva) (Figura 7), o puede estar en otras superficies que fueron contaminadas

por personas infectadas con el virus. El SARS-CoV-2 es un virus de RNA con nucleoproteínas en su interior, la cápsula está formada por una capa de lípidos, razón por la cual al lavarse adecuadamente con jabón o alcohol, la capa lipídica se desordena provocando que el virus se desarme, perdiendo toda función infectiva. Hay varias glicoproteínas (proteínas “adornadas” con azúcares), en la membrana viral entre las que destaca la proteína trimérica llamada *spike* o espícula [27]. Ahora sabemos que si esta proteína tiene la conformación abierta en su dominio de unión al receptor o RBD (dominio receptor obligatorio, del inglés: *receptor binding domain*; localizado entre los residuos 331-524) (Figura 8) [28], se inicia el mecanismo infectivo del virus. Así se reconoce al receptor de angiotensina 2 presente en las células humanas, para desencadenar la multipli-

Gotita de saliva

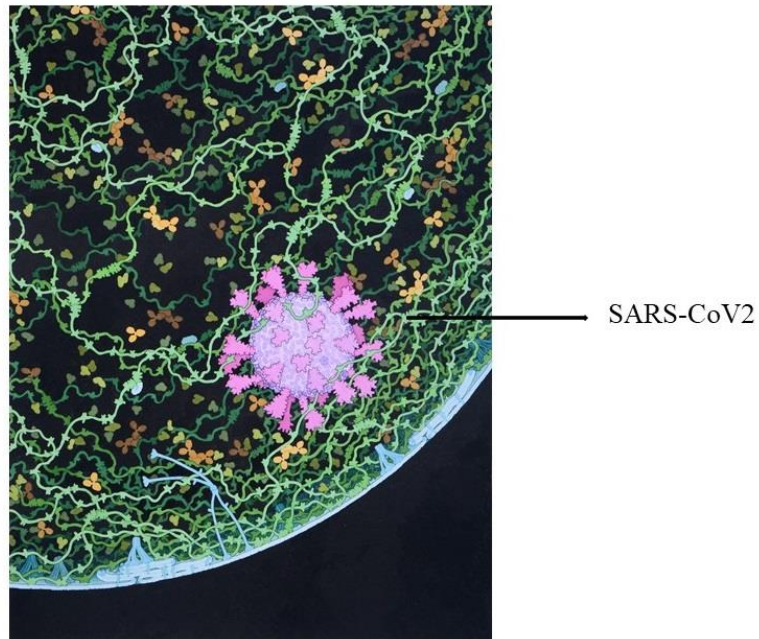
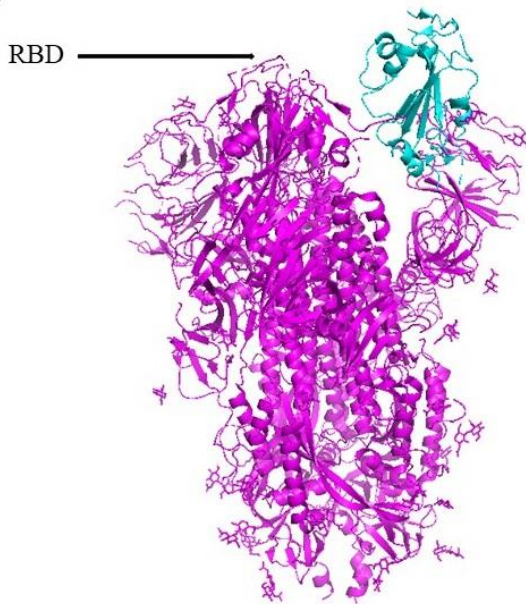


Figura 7. Pintura basada en evidencia experimental sobre el SARS-CoV2, el coronavirus (Lila) y las proteínas de la corona entre las que se encuentra la espícula o spike (rosa). Figura tomada de David S. Goodsell. RCSB Protein Data Bank; doi: 10.2210/rcsb_pdb/goodsell-gallery-024.

A)



B)

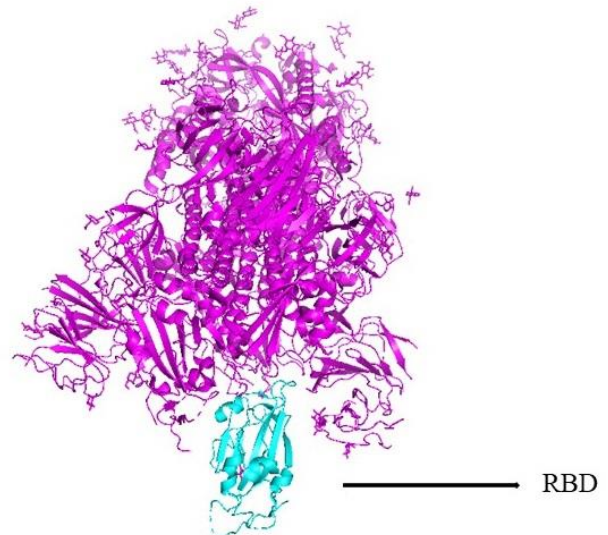


Figura 8. A) Estructura tridimensional del monómero de la proteína de la espícula del SARS-Cov2 en su estado abierto en color fiusha y dominio RBD en color cian código de PDB (6VSB) vista lateral y B) vista aérea determinada por Crio-Microscopia electrónica vistas en PyMOL, en esta conformación el virus es muy infeccioso. Figuras realizadas con la información de <https://www.rcsb.org/structure/6VSB> y modificadas en PyMOL.

cación de más virus que serán expulsados para volver a iniciar el ciclo de virulencia en las células del enfermo, o en otros individuos. El uso de esta proteína espícula como inmunógeno para la generación de vacunas (Figura 9) [29], induce la producción de anticuerpos neutralizantes de la acción del virus SARS-CoV-2 y potencialmente, el fin de la pandemia si alcanzamos tasas de vacunación altas en nuestras comunidades. Si revisamos los códigos del PDB: 6VXX y 6VSB, los cuales corresponden a estructuras de la espícula en su conformación cerrada [30] y abierta [28] determinadas ambas por criomicroscopía electrónica. Al ingresar el código 6VXX en el PDB aparece la información relacionada a la estructura tridimensional de la glicoproteína de la espícula del SARS-Cov-2 en su estado cerrado (Figura 9), además de más información acerca de la estructura en la siguiente liga <https://www.rcsb.org/structure/6VXX>.

Si revisamos en el PDB el código de cuatro letras 6VSB encontraremos la estructura tridimensional de la conformación abierta del dominio RBD (Figura 8) y más información sobre las características de esta estructura en la siguiente [liga](https://www.rcsb.org/structure/6VSB) <https://www.rcsb.org/structure/6VSB>.

Estos son dos ejemplos de lo que podemos encontrar como usuarios en dos estructuras determinadas por la criomicroscopía electrónica, teniendo contribuciones muy importantes sobre el entendimiento de la proteína espícula involucrada en la infección del virus SARS-Cov-2 sobre células humanas. Todo lo descrito en este artículo muestra la

trascendencia del PDB mediante el análisis de tres interesantes ejemplos de estructuras determinadas por las técnicas de difracción de rayos X (Proteína GFP de la medusa *Aequorea victoria*), RMN (toxina Cn2 del alacrán *C. noxius*) y criomicroscopía electrónica (proteína de la espícula del coronavirus SARS-CoV-2) (Figura 10). En cada una de ellas se obtiene información muy valiosa para diferentes áreas de investigación. La integración de estas diferentes técnicas, permitirá entender mejor la biología estructural de las macromoléculas e incluso comprender y atender procesos que ahora no podemos.

CONCLUSIÓN

El *Protein Data Bank* es una base de datos de estructuras muy grande y es a la vez fuente de mucha información importante sobre diversas macromoléculas, las cuales pueden ser analizadas por una diversidad de profesionales de diversas áreas como cristalógrafos, químicos, biólogos, bioquímicos, microbiólogos, médicos, y farmacéuticos. El conocer la estructura de una macromolécula de interés puede ayudar a entender una gran cantidad de enfermedades como las enfermedades respiratorias provocadas por diversos virus, las enfermedades intestinales como la salmonelosis, u otros padecimientos médicos más complejos como el Parkinson, el Alzheimer, o la diabetes, ya que conociendo a detalle atómico la alteración química y estructural de una macromolécula podemos descifrar los mecanismos que provocan la enfermedad y el daño en el organismo; además

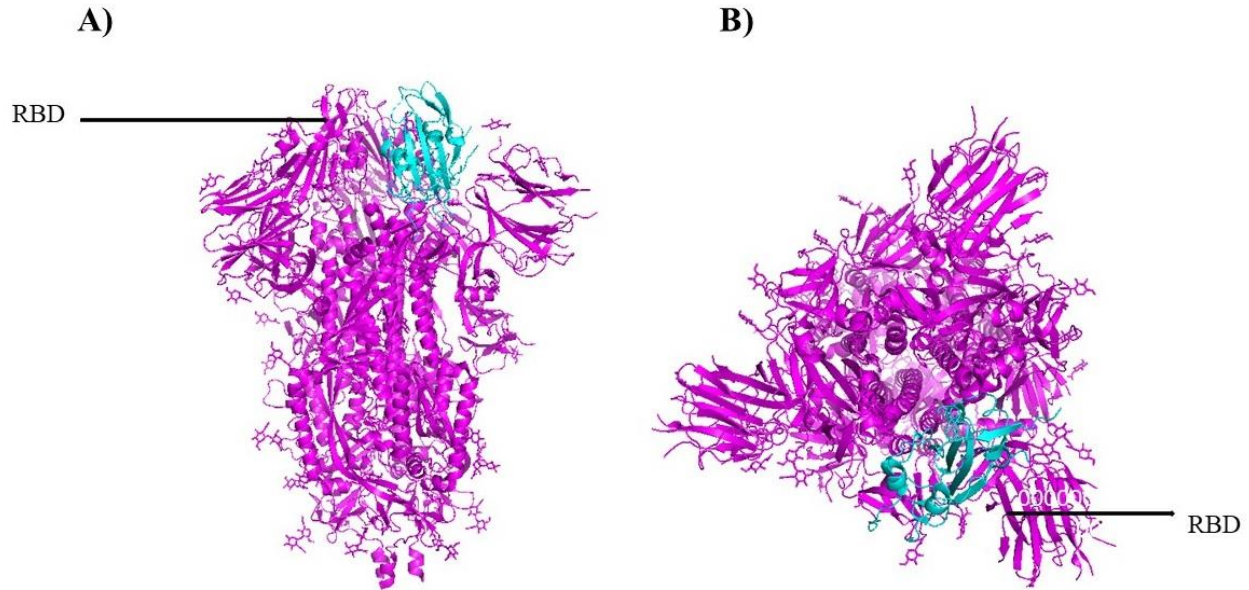


Figura 9. A) Estructura tridimensional del monómero de la proteína de la espícula del SARS-CoV2 en su estado cerrado en color fiusha y dominio RBD en color cian código de PDB (6VXX) vista lateral y B) vista aérea determinada por Crio-Microscopia electrónica vistas en PyMOL, en esta conformación el virus no es infeccioso. Figuras realizadas con la información de <https://www.rcsb.org/structure/6VXX> y modificadas en PyMOL.

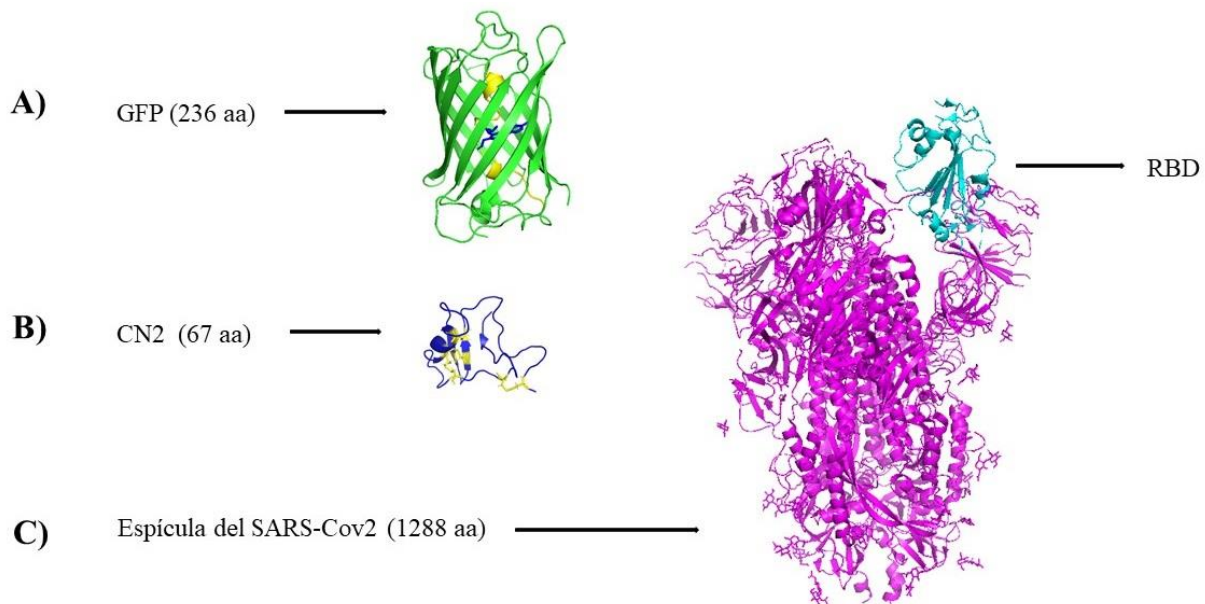


Figura 10. Estructuras tridimensionales de los 3 ejemplos reportados en este artículo: A) GFP PDB (1EMA), determinada por Difracción de Rayos X, B) Toxina Cn2 PDB: (1CN2) determinada por NMR y proteína monomérica de la espícula en su estado abierto PDB (6VSB) determinada, por Crio-Microscopia electrónica. Visualizadas con PyMol. El número de aminoácidos de cada proteína se indica entre paréntesis.

se puede realizar el diseño de fármacos para su tratamiento o incluso para la cura de algunos padecimientos.

Este es el link para consultar y descargar el calendario del año 2021 que celebra el 50 aniversario del PDB:
<https://pdb101.rcsb.org/learn/flyers-posters-and-other-resources/calendar/2021-calendar-celebrating-50-years-of-protein-data-bank>

CONFLICTO DE INTERESES

Los autores declaran la ausencia de conflicto de interés.

AGRADECIMIENTOS

Los autores agradecen al Mtro. Juan Manuel Hurtado Ramírez por su asistencia técnica en computación.

REFERENCIAS

- [1]. Berman HM, Battistuz T, Bhat TN, Bluhm WF, Bourne PE, Burkhardt K, *et al.* The protein data bank. *Acta Crystallogr Sect D Biol Crystallogr.* 2002;58(6 I):899–907.
- [2]. Prestegard JH. REVIEWS A perspective on the PDB 's impact on the field of glycobiology. 2021;(8):1–11.
- [3]. Chadarevian S De. John Kendrew and myoglobin : Protein structure determination in the 1950s. 2018;27:1136–43.
- [4]. Li Z, Albert J, Ve C. topical reviews Cryo-electron microscopy and X-ray crystallography : complementary approaches to structural biology and drug discovery 1.

2017;174–83.

- [5]. Ferry G. Osamu Shimomura (1928–2018). *Nature.* 2018;563(7733):627–627.

- [6]. Günther S, Reinke PYA, Fernández-García Y, Lieske J, Lane TJ, Ginn HM, *et al.* X-ray screening identifies active site and allosteric inhibitors of SARS-CoV-2 main protease. 2021;646(May):642–6.

- [7]. Ishima R, Torchia DA. Protein dynamics from NMR. *Nat Struct Biol.* 2000;7(9):740–3.

- [8]. Rhs CJ. Magnetic resonance spectroscopy. 2017;100:241–92.

- [9]. Murata K, Wolf M. Cryo-electron microscopy for structural analysis of dynamic biological macromolecules. *Biochim Biophys Acta - Gen Subj* [Internet]. 2018;1862(2):324–34. Available from: <https://doi.org/10.1016/j.bbagen.2017.07.020>

- [10]. Carroni M, Saibil HR. Cryo electron microscopy to determine the structure of macromolecular complexes. *Methods* [Internet]. 2016;95:78–85. Available from: <http://dx.doi.org/10.1016/j.ymeth.2015.11.023>

- [11]. Hori Y. Crystal structure of the *Aequorea victoria* green fluorescent protein. *Tanpakushitsu Kakusan Koso.* 2007;52(13 Suppl):1768–9.

- [12]. Gaytán P, Roldán-Salgado A, Yáñez JA, Morales-Arrieta S, Juárez-González VR. CiPerGenesis, A. Mutagenesis Approach that Produces Small Libraries of Circularly Permuted Proteins Randomly Opened at a Focused Region: Testing on the Green Fluorescent Protein. *ACS Comb Sci.*

2018;20(7):400–13.

[13]. Sawano A, Miyawaki A. Directed evolution of green fluorescent protein by a new versatile PCR strategy for site-directed and semi-random mutagenesis. *Nucleic Acids Res.* 2000;28(16).

[14]. Chalfie M. GFP: lighting up life (Nobel Lecture). *Angew Chem Int Ed Engl.* 2009;48(31):5603–11.

[15]. Cressey D. Roger Tsien's legacy: The creations that lit up biology. *Nature.* 2020;(September):1–5.

[16]. Remington SJ. Green fluorescent protein: A perspective. *Protein Sci.* 2011;20(9):1509–19.

[17]. Yang F, Moss LG, Phillips GN. The Molecular Structure of Green Fluorescent Protein. *Nat Biotechnol.* 1996;14(10):1246–51.

[18]. Zimmer M. Green fluorescent protein (GFP): Applications, structure, and related photophysical behavior. *Chem Rev.* 2002;102(3):759–81.

[19]. Ormo M, Cubitt AB, Kallio K, Gross LA, Tsien RY, Remington SJ. Crystal Structure of the *Aequorea victoria* Green Fluorescent Protein Author (s): Mats Ormö , Andrew B . Cubitt , Karen Kallio , Larry A . Gross , Roger Y . Tsien and Published by: American Association for the Advancement of Science Stable URL : [http://. Science \(80- \)](http://. Science (80-)). 1996;273(5280):1392–5.

[20]. Selisko B, Licea AF, Becerril B, Zamudio F, Possani LD, Horjales E. Antibody BCF2 Against Scorpion Toxin Cn2 From

Centruroides noxius Hoffmann: Primary Structure and Three-Dimensional Model as Free Fv Fragment and Complexed With Its Antigen. 2001;143(January 1999):130–43.

[21]. Juárez-González VR, Riaño-Umbarila L, Quintero-Hernández V, Olamendi-Portugal T, Ortiz-León M, Ortiz E, *et al.* Directed evolution, phage display and combination of evolved mutants: A strategy to recover the neutralization properties of the scFv version of BCF2 a neutralizing monoclonal antibody specific to scorpion toxin Cn2. *J Mol Biol.* 2005;346(5):1287–97.

[22]. Quintero-Hernández V, Juárez-González VR, Ortiz-León M, Sánchez R, Possani LD, Becerril B. The change of the scFv into the Fab format improves the stability and *in vivo* toxin neutralization capacity of recombinant antibodies. *Mol Immunol.* 2007;44(6):1307–15.

[23]. Pintar A, Possani LD, Delepierre M. Solution structure of toxin 2 from *Centruroides noxius* Hoffmann, a β -scorpion neurotoxin acting on sodium channels. *J Mol Biol.* 1999;287(2):359–67.

[24]. Canul-tec JC, Rian L, Rudin E, Torres-larios A. Structural Basis of Neutralization of the Major Toxic Component from the Scorpion *Centruroides noxius* Hoffmann by a Human-derived Single-chain Antibody Fragment. 2011;286(23):20892–900.

[25]. Rian L, Ledezma-candanoza LM, Serrano-Posada H, Ferna G, Olamendi-portugal T, Rojas-trejo S, *et al.* Optimal Neutralization of *Centruroides noxius* Venom Is Understood

through a Structural Complex between Two Antibody Fragments and the Cn2 Toxin. 2016;291(4):1619–30.

[26]. Elizalde González JJ. SARS-CoV-2 y COVID-19. Una revisión de la pandemia. *Med Crítica*. 2020;33(1):53–67.

[27]. Casalino L, Gaieb Z, Dommer AC, Harbison AM, Fogarty CA, Barros EP, *et al.* Shielding and Beyond: The Roles of Glycans in SARS-CoV-2 Spike Protein. 2020; 28;6(10):1722-1734.

[28]. Wrapp D, Wang N, Corbett KS,

Goldsmith JA, Hsieh CL, Abiona O, et al. Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science* (80-). 2020;367(6483):1260–3.

[29]. Wilson IA. 50 Years of structural immunology. 2021;1–20.

[30]. Walls AC, Park YJ, Tortorici MA, Wall A, McGuire AT, Veesler D. Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. *Cell* [Internet]. 2020;181(2):281-292.e6. Available from: <http://dx.doi.org/10.1016/j.cell.2020.02.058>