CÁNCER: LAS CÉLULAS QUE NO NECESITAN OXÍGENO PARA SOBREVIVIR

CANCER: CELLS THAT DO NOT NEED OXYGEN TO SURVIVE

Osvaldo Jimenez-Montiel (1) Norma A. Caballero (2) Francisco J. Meléndez (3) María Eugenia Castro (4)

> https://orcid.org/0009-0001-1609-7949 https://orcid.org/0000-0002-9410-5852 https://orcid.org/0000-0002-5796-0649 https://orcid.org/0000-0003-1716-7707

> > Año 10 No. 30 Recibido: 10/marzo/2024 Aprobado: 15/mayo/2024 Publicado: 10/septiembre/2024

(1) Estudiante de Maestría en Ciencias Químicas,
Facultad de Ciencias Químicas, BUAP
(2) Facultad de Ciencias Biológicas, BUAP, 72570, México
(3) Laboratorio de Química Teórica, Facultad
de Ciencias Químicas BUAP, 72570, Puebla, México
(4) Centro de Química del Instituto de Ciencias,
ICUAP, BUAP, 72570, Puebla, México
(222) 2295500 ext. 2830 y 2819
osvaldo.jimenezmo@alumno.buap.mx*
norma.caballero@correo.buap.mx
francisco.melendez@correo.buap.mx
mareug.castro@correo.buap.mx

Resumen

El oxígeno es el principal componente que requieren las células de un organismo vivo para obtener energía y de esta manera sea capaz de realizar sus actividades metabólicas de forma adecuada. Entonces, ¿podrán existir algunas células sin el suministro de oxígeno? La respuesta es sí. Por increíble que parezca, existen células en diversos padecimientos que son capaces de permanecer vivas a bajas concentraciones de oxígeno. Esto es posible debido a una serie de mecanismos moleculares que se activan al percibir este estado de hipoxia, como la expresión de la familia del factor inducido por hipoxia (HIF-1, 2 y 3), que protegen a las células en padecimientos como la anemia, la isquemia miocárdica, la hipertensión pulmonar, el cáncer, etc. Específicamente, la familia HIF brinda resistencia a las células presentes en cánceres sólidos contra las terapias actuales y aseguran la supervivencia del tumor cambiando de su metabolismo aerobio a un metabolismo anaerobio, generando nuevas vascularizaciones para una mayor irrigación sanguínea y activando genes encargados de aumentar la producción de células sanguíneas dentro del organismo, etc. Debido al papel tan importante que desempeña la familia HIF en la supervivencia de células cancerígenas, en esta investigación se propone describir el bloqueo de la actividad de HIF-1 empleando metodologías in silico. Se considera que un punto clave para su inactivación es evitar la dimerización de las subunidades proteicas que lo constituyen. Como resultado de este análisis, se pretende tener un nuevo aliado en las terapias actuales contra el cáncer.

Palabras clave: Cáncer, Factor HIF-1, Dimerización, Química computacional, Acoplamiento molecular.

Abstract

Oxygen is the main component required by the cells of a living organism to obtain energy and thus be able to perform its metabolic activities properly. So, can some cells exist without the supply of oxygen? The answer is yes. Incredible as it may seem, there are cells in various diseases that can remain alive at low oxygen concentrations. This is possible due to a series of molecular mechanisms that are activated upon perceiving this state of hypoxia, such as the expression of the hypoxiainduced factor family (HIF-1, 2 and 3), which protect cells in conditions such as anemia, myocardial ischemia, pulmonary hypertension, cancer, etc. Specifically, the HIF family provides resistance to the cells present in solid cancers against current therapies and ensures the survival of the tumor by changing its aerobic metabolism to an anaerobic metabolism. generating new vascularization for greater blood irrigation and activating genes responsible for increasing the production of blood cells within the organism, etc. Due to the significant role played by the HIF family in the survival of cancer cells, we propose to describe the blocking of HIF-1 activity using in silico methodologies. It is considered that a key factor for its inactivation is the non-dimerization of its constituent protein subunits. As a result of this analysis, it is intended to have a new ally in the current therapies against cancer.

Keywords: Cancer, HIF-1 factor, Dimerization, Computational chemistry, Molecular docking.

¿Cómo sobrevivir sin oxígeno?

El oxígeno se requiere para el correcto funcionamiento de organismos con metabolismo aerobio, manteniendo la demanda de los gastos bioenergéticos de las células y los tejidos. Si la concentración de oxígeno disminuye considerablemente, las consecuencias pueden ser fatales, ya que morirían los tejidos comprometidos por estos bajos niveles de oxígeno (Bertout et al., 2008). La isquemia miocárdica es un padecimiento representativo de esta condición. la cual consiste en la disminución del diámetro de las arterias coronarias del corazón v reduce como consecuencia el flujo sanguíneo, por lo tanto, el músculo cardíaco no recibe suficiente oxígeno. Otro padecimiento es la generación de trombos cerebrales, en donde la pérdida de oxígeno se debe a la obstrucción de la vascularización por un coágulo sanguíneo. En procesos como el cáncer hay una proliferación celular descontrolada cuya agregación forma cúmulos densos (cáncer de tipo sólido) cuvas células más profundas carecen de oxígeno (Chao et al., 2021). ¿Cómo sobreviven este tipo de células que no tienen oxígeno? Existen tres sistemas que intervienen en la respuesta a la hipoxia: a) de detección o sensor de oxígeno; b) de regulación, mediante el control de la expresión de una amplia serie de genes; y c) efector múltiple, incluye no solo expresión de genes, sino múltiples cambios funcionales que van desde la estimulación de moléculas vasodilatadoras hasta las variaciones en la afinidad de la hemoglobina por el oxígeno. El sistema regulador está modulado directamente por el sensor v el elemento organizador principal es un factor de transcripción específico, el factor inducible por hipoxia 1 (HIF-1). Este factor pertenece a una familia de proteínas compuesta por tres elementos (HIF-1, HIF-2 y HIF-3) que son activados ante las bajas concentraciones de oxígeno en el tejido, pero trabajan de manera distinta. En esta ocasión, nos centraremos específicamente en HIF-1 que es el elemento más estudiado.

Estructuralmente, HIF-1 es un factor de transcripción proteico compuesto por dos

partes llamadas subunidades, la subunidad (v la subunidad ®. Αμβασ estructuras cuentan con regiones específicas que les permiten unirse una a la otra para conformar el factor de transcripción que le hará frente a la falta de oxígeno donde sea necesario (Figura 1). La subunidad (se transcribe de forma constitutiva, es decir. se encuentra presente de forma continua al interior de la célula, pero si esto es así, ¿cómo es que no está activa en todo momento? La respuesta reside en que esta subunidad es sensible a la presencia de oxígeno de tal manera que es "marcada" por otras proteínas para ser degradada, es decir, si las concentraciones de oxígeno son óptimas, el HIF-1 (debe ser eliminado. ¿Pero qué pasa cuando no hay oxígeno en el medio celular? Es aquí cuando HIF-1(no es degradado y comienza el proceso de respuesta ante la condición de hipoxia uniéndose a la subunidad ® formando el heterodímero HIF-1(/®, el cual viajará al núcleo y transactivará genes que permitirán la supervivencia de las células. Estos genes están implicados en el cambio de un metabolismo aerobio a uno anaerobio, en la generación de nuevos vasos sanguíneos para suplir la ausencia de oxígeno debido a una isquemia. la activación de señales que permitan la formación de más células sanguíneas para suministrar una mayor cantidad de oxígeno, e incluso la sobre expresión de transportadores de glucosa (Brown, 2007).

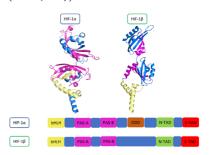


Figura 1. Estructura 3D del complejo proteico HIF-1. En magenta se resaltan los sitios de unión PAS-A Y PAS-B necesarios para la dimerización de la subunidad ((izquierda) y ® (derecha). En amarillo se muestra la región de unión a ADN, bHLH.

En el cáncer, donde existe un crecimiento celular descontrolado, las células más internas se ven ausentes de irrigación sanguínea, lo que significa que tales células quedan desprovistas de oxígeno

y necesitan de la activación de los factores de transcripción HIF para no morir (Schönberger et al., 2021).

Un nuevo aliado para las terapias existentes

La familia de transcripción HIF realiza funciones que se contraponen, por un lado, ayuda a enfrentar deficiencias de oxígeno como en desórdenes del corazón, desórdenes cerebrovasculares, anemias, leucemias y procesos de cicatrización: y, por otro lado, proporciona resistencia a las células cancerígenas de tumores sólidos (sarcomas y carcinomas) que se encuentran en condiciones extremas de deficiencia de oxígeno. Desde la década de los 50's, Gray y col. mostraron que las células no tumorales presentes en condiciones de baia concentración de oxígeno son protegidas ante los efectos citotóxicos que pudieran sufrir al ser expuestas a rayos-X (Gray et al., 1953). Posteriormente, a principios del siglo XXI, Brown describe que los tumores sólidos humanos están invariablemente menos oxigenados que los teiidos normales de los que provienen, y que esta baja concentración de oxígeno provoca que las células tumorales presenten resistencia a la radioterapia y quimioterapia (Brown, 2007). Por lo tanto, HIF es una herramienta de supervivencia de células malignas y un escudo protector contra las terapias que buscan destruir dichas células.

En años más recientes, Minassian y col., aseveran que la resistencia de los tumores malignos a la quimioterapia se debe al microambiente deficiente de oxígeno y que esta condición puede ser un factor significativo en el desarrollo de la resistencia a fármacos (Minassian et al., 2019). Con base en lo anterior, se ha planteado que la inhibición de al menos un miembro de la familia de transcripción HIF sería un gran complemento para las terapias ya existentes.

Se ha visto que la desregulación de la vía de HIF-1 es clave para detener las características malignas de un tumor, como el crecimiento descontrolado y la generación de nuevas vascularizaciones. Esta vía comienza con la expresión del ARN mensajero de HIF que llevará a la síntesis de las subunidades (y ®, posteriormente ambas proteínas maduran y se unen entre sí para generar el dímero HIF-1(/®. La unión de este dímero con otras proteínas formarán un complejo de transcripción y, por último, se expresarán los genes que son objetivo de HIF-1(/®.

Con el conocimiento de esta vía, varios grupos de investigación enfocados en cáncer se han dado a la tarea de proponer compuestos químicos que la bloqueen en las diferentes etapas que la constituyen. Desde evitar la expresión de HIF-1 a partir del ADN hasta impedir que HIF acompleiado transcriba las regiones de ADN que permitirán la supervivencia de las células tumorales. Por ejemplo, Vorinostat (SAHA) es un compuesto que bloquea la expresión del ARN mensajero. Bortezomib (Velcade®) inhibe la síntesis de la subestructura (, oX3 se ocupa de impedir la unión de las subestructuras (y ®, acriflavina actúa en la inhibición del compleio de transcripción de las subestructuras (y ® con otros elementos llamados CPB y p300, y Everolimus, un bloqueador de las regiones de ADN donde HIF tiene la actividad de transcripción (Minassian et al., 2019).

La secuencia de la vía de HIF-1

HIF-1 realiza la activación de los genes que ayudarán a las células cancerígenas a sobrevivir, pero ¿cómo lo hace? Mediante un proceso en el que es necesario la unión inicial de dos proteínas. Entonces, si HIF-1 funciona mediante dos proteínas, se sabe de antemano que para que existan debe haber un gen o un grupo de genes que por ARN mensaiero en el ribosoma formarán a esas proteínas inmaduras. Lo que posteriormente madurarán por procesos de glicosilación, se fusionarán, formarán el complejo HIF-1(/® que una vez en el núcleo activará a los genes que necesita el tumor maligno. En la Figura 2 se muestran seis puntos bien definidos de la vía en los que se han planteado potenciales fármacos dirigidos a interrumpir la actividad de HIF-1 (Schönberger et al., 2021).

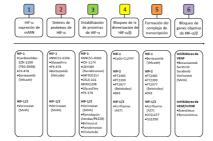


Figura 2. Puntos definidos de la vía de HIF-1 en donde se han probado compuestos químicos específicos que conllevan a su desregulación. Modificada de (Schönberger et al., 2019)

Simulación computacional como alternativa en la investigación

Los métodos computacionales de simulación molecular se han convertido en una herramienta de suma importancia en el estudio de distintas enfermedades para poder combatirlas. La experimentación in sillico trae consigo bastantes ventaias. por ejemplo, facilita modelar estructuras químicas que funcionen como inhibidores que, si lo extrapolamos al cáncer, serían aquellas moléculas capaces de bloquear la secuencia de la vía en cualquiera de los seis puntos mencionados en la sección anterior (ver Figura 2). Por otro lado, y a pesar de que las metodologías experimentales han avanzado a pasos agigantados, existen moléculas como las proteínas que no se han podido elucidar estructuralmente, por lo que el modelarlas resulta ser la mejor opción y predecir su comportamiento al colocarlas en condiciones específicas.

El acoplamiento molecular, también conocido como docking molecular, es una
herramienta computacional que permite
analizar las interacciones intermoleculares que se puedan llevar a cabo entre un
ligante (molécula pequeña) y un receptor
(macromolécula), además de predecir la
estabilidad del complejo resultante. El
tiempo invertido en esta experimentación
in silico es mucho menor que el que se necesitaría para llevar a cabo el experimento
in vitro. Por otro lado, el ahorro económico

también es un factor que se debe resaltar, ya que la inversión monetaria en la compra de reactivos o bien de insumos necesarios para su síntesis, así como el aislamiento, purificación y cristalización de proteínas in vitro es sumamente alta. Cabe mencionar que el docking molecular es la primera técnica de elección empleada en el diseño de fármacos (Meng et al., 2011)

Modelado de inhibidores

Antes de llevar a cabo un acoplamiento molecular, deben modelarse de manera independiente al ligante y al receptor debido a las diferentes propiedades químicas y estructurales de cada uno. Los datos experimentales siempre serán el punto de partida ideal para una buena simulación. por lo que se buscan en las diversas bases de datos cristalográficas, aunque no siempre se tiene éxito. En la actualidad se han reportado algunos compuestos químicos con actividad anticancerígena que actúan inhibiendo la actividad de HIF-1 bloqueando la dimerización de las subunidades (v ®. Uno de estos compuestos es el llamado oX3, que es un derivado de hidroxibenzaldehído, y a pesar de que se ha reportado que tiene actividad inhibitoria en HIF-2, se contempla como inhibidor en HIF-1 por ser una proteína homóloga de HIF-2.

Por otro lado, los compuestos PT2377 (Belzitufán), PT2399 y PT2385, que son derivados de inden-fluorobenzonitrilo, inhibidores de HIF-2 que, por las mismas razones que para oX3, pueden ser potenciales inhibidores de HIF-1.

Estos inhibidores (Figura 3) se modelaron en el programa Gaussian16 (Frisch et al., 2016) para obtener las estructuras de mínima energía, lo que, en otras palabras, significa, obtener la estructura molecular que en ciertas condiciones será la más estable. Se utilizó la teoría del funcional de la densidad (DFT) (Hohenberg & Kohn, 1964) usando el nivel de teoría B 3LYP/6-31+G(d) (Becke,1993; Ditchfield et al., 1971).

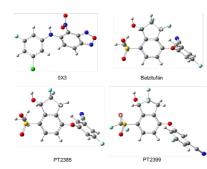


Figura 3. Estructuras 3D de los ligantes bloqueadores de la unión de las subunidades (y ® de HIF. Los átomos se presentan en código de color CPK.

Modelado de la macromolécula

Las bases de datos biológicos disponibles para obtener información de proteínas resueltas cristalográficamente son bastante amplias, sin embargo, no se encuentran todas las proteínas que existen en la naturaleza. Se realizó la búsqueda en la base de datos Protein Data Bank (https:// www.rcsb.org) de la estructura cristalina de HIF-1(/® humana, sin embargo, no se encontraron resultados, por lo que fue necesario utilizar técnicas de modelado de proteínas. El modelado de la proteína se realizó mediante un alineamiento múltiple de secuencias de aminoácidos (AA) de la proteína que se desea modelar con una aserie de secuencia de aminoácidos de proteínas homólogas con estructuras cristalinas resueltas (Figura 4). Para dicho alineamiento se utilizó el programa Clustal Omega (Thompson et al., 1994) y se visualizó en el programa JalView (Waterhouse et al., 2009).

Una vez alineadas las secuencias se lleva a cabo el modelado por homología con el programa Modeller (Webb & Sali, 2016), que con base en las estructuras cristalinas y el resultado del alineamiento de secuencias realizará la construcción de la estructura proteica deseada.

Se utilizaron tres estructuras como molde para llevar a cabo al modelado de la estructura de HIF-1〈 con PDB ID 4ZP4, 4ZPR, 6E3T pertenecientes a la estructura HIF-2〈, HIF-1〈, y HIF-2〈/® aisladas de ratón (Mus musculus), respectivamente. Además de realizar un modelo de la proteína a partir de las estructuras molde, se realizó un solo modelo utilizando los tres moldes, al cual se denomina modelado múltiple. Las estructuras resultantes se muestran en la Figura 5.

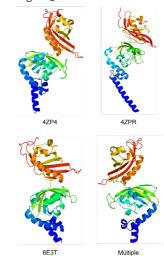


Figura 5. Estructuras 3D de las estructuras modeladas de la subunidad (de HIF-1 de humano a partir de los diferentes moldes utilizados.

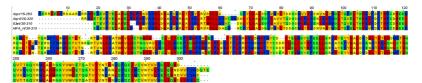


Figura 4. Alineamiento por homología para modelado de proteínas. La fila uno corresponde a la secuencia de aminoácidos (AA) de la proteína a modelar y las filas subsecuentes corresponden a la secuencia de AA de proteínas homólogas cuyas estructuras cristalinas se encuentran resueltas.

Modelado de la interacción macromolécula-inhibidor

Una vez obtenidas las estructuras de los inhibidores y de las proteínas (modeladas, el siguiente paso a realizar es el acoplamiento entre cada subunidad modelada a partir de los diferentes moldes y los inhibidores (Figura 6), con el objetivo de encontrar información relevante para la descripción de un posible mecanismo por el cual estos inhibidores obtienen su actividad anticancerígena.



Figura 6. Diagrama general de trabajo para el acoplamiento molecular inhibidor-HIF-1α modelada.

De los cálculos de acoplamiento molecular realizados con el programa Autodock Vina (Trott & Olson, 2010; Eberhardt et al., 2021) se obtuvo información relevante como las energías de afinidad, siendo el valor más bajo para el complejo formado por PT2385-HIF-1((4ZP4) de -8.7 kcal/mol. El sitio de interacción se identificó en el dominio

de unión PAS-A, por lo que el inhibidor impediría la unión de esta subunidad (con la correspondiente subunidad (Figura 7). Las interacciones presentadas para lograr su estabilidad fueron bastante diversas, desde interacciones débiles de van der Waals hasta interacciones de tipo puente de hidrógeno. Es importante resaltar que el átomo de flúor del inhibidor muestra un papel relevante al formar puentes de hidrógeno de tipo halógeno (Figura 8).

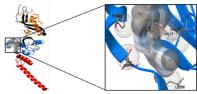


Figura 7. Representación gráfica del acoplamiento molecular de PT2358 con la proteína HIF-1α de humano modelada.

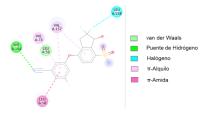


Figura 8. Diagrama 2D del complejo PT2385-HIF-1 α . Los diferentes tipos de interacción se representan en código de colores. Las interacciones de tipo π -alquilo entre la Valina y los anillos aromáticos del inhibidor son predominantes.

Perspectivas

El estudio computacional de las vías de los factores de transcripción HIF representa un campo de investigación muy amplio, ya que el papel que desempeña cada uno de ellos (HIF-1, HIF-2 y HIF-3) en presencia de masas tumorales es distinto y su expresión depende del tipo de cáncer.

Hasta el momento se tienen los indicios del posible mecanismo de acción de algunos inhibidores propuestos que impiden la unión de las cadenas < y ® de HIF-1, siendo el dominio de unión PAS-A y la presencia de átomos de flúor relevantes para esta inhibición. La continuación de este trabajo consiste en seguir explorando la vía de HIF con los diversos inhibidores experimentalmente reportados que actúan en diferentes pasos de la vía, de tal forma que en un futuro no muy lejano se sugiera una estructura química, basada en moléculas de origen natural, más efectiva en el tratamiento de cáncer sólido.

Por último, se incentiva a continuar aportando a la ciencia con cualquier indicio que ayude a combatir enfermedades tan terribles como el cáncer.

Declaración de no Conflicto de intereses

Los autores declaran que no existe conflicto de interés alguno.

Declaración de privacidad

Los datos de este artículo, así como los detalles técnicos para la realización del experimento, se pueden compartir a solicitud directa con el autor de correspondencia.

Los datos personales facilitados por los autores a RD-ICUAP se usarán exclusivamente para los fines declarados por la misma, no estando disponibles para ningún otro propósito ni proporcionados a terceros.

Agradecimientos

O.J.M. agradece al Consejo Nacional de Ciencia y Tecnología (CONACYT) por la beca de Maestría (número 1173751). Los autores agradecen a la Vicerrectoría de Investigación y Estudios de Posgrado (VIEP-BUAP), al Laboratorio Nacional de Supercómputo del Sureste de México (LNS-BUAP) por los recursos computacionales y al Cuerpo Académico BUAP-CA-263 de PRODEP (SEP, México).

Referencias

- Becke, A. D. (1993). Density-functional thermochemistry. III. The role of exact exchange, Journal of Chemical Physics, 98 5648-52. doi:10.1063/1.464913
- Bertout, J. A., Patel, S. A., & Simon, M. C. (2008). The impact of O2 availability on human cancer. Nature Reviews Cancer, 8(12), 967-975. doi:10.1038/nrc2540
- Brown, J. M. (2007). Tumor hypoxia in cancer therapy. Methods in enzymology, 435, 295-321. doi:10.1016/S0076-6879(07)35015-5
- Chao, Y. A. N. G., Zhong, Z. F., Sheng-Peng, W. A. N. G., Chi-Teng, V. O. N. G., Bin, Y. U., & Yi-Tao, W. A. N. G. (2021). HIF-1: Structure, biology, and natural modulators. Chinese Journal of Natural Medicines, 19(7), 521-527. doi:10.1016/S1875-5364(21)60051-1
- Ditchfield, R., Hehre, W. J., & Pople J. A. (1971). Self-consistent molecular-orbital methods. 9. Extended Gaussian-type basis for molecular-orbital studies of organic molecules. Journal of Chemical Physics, 54(2), 724-728. doi:10.1063/1.1674902
- Eberhardt, J., Santos-Martins, D., Tilack, A. F., Forli, S. (2021). AutoDock Vina 1.2.0: New docking methods, expanded force field, and python bindings. Journal of Chemical Information and Modeling, 61, 8, 3891-3898. doi:10.1021/acs.jcim.1coo203
- Frisch, M. I., Trucks, G. W., Schlegel, H. B., et al., Gaussian 16. Revision, B.o1 2016; Gaussian Inc.: Pittsburgh, PA, USA, 2016.
- Gray, L. H., Conger, A., Ebert, M., Hornsey, S., & Scott, O. C. A. (1953). The concentration of oxygen dissolved in tissues at the time of irradiation as a factor in radiotherapy. The British Journal of Radiology, 26(312), 638-648. doi:10.1259/0007-1285-26-312-638
- Hohenberg, H., & Kohn, W., (1964) Inhomogeneous Electron Gas, Physical Review, 136, B864-B871. doi:10.1103/PhysRev.136.B864
- Minassian, L. M., Cotechini, T., Huitema, E., & Graham, C. H. (2019). Hypoxia-induced resistance to chemotherapy in cancer. In Hypoxia and Cancer Metastasis (pp. 123-139). Springer, Cham. doi:10.1007/978-3-030-12734-3_9
- Meng X. Y., Zhang H. X., Mezei M., & Cui M. Molecular docking: a powerful approach for structure-based drug Discovery (2011). Current Computer-Aided Drug Design, 7(2):146-57. doi:10.2174/157340911795677602.
- Schönberger, T., Fandrey, J., & Prost-Fingerle, K. (2021). Ways into understanding HIF inhibition. Cancers, 13(1), 159. doi:10.3390/ cancers13010159

- Thompson, J. D., Higgins, D. G., & Gibson, T. J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Research, 22(22), 4673-4680. doi:10.1093/nar/22.22.4673
- Trott, O. & Olson, A. J. (2010) AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. Journal of Computational Chemistry, 31(2), 455-461. doi:10.1002/jcc.21334
- Waterhouse A. M., Procter J. B., Martin D. M. A., Clamp M., & Barton G. J. (2009) Jalview Version 2 - A multiple sequence alignment editor and analysis workbench. Bioinformatics 25 1189-1191. doi:10.1093/ bioinformatics/btpo33
- Webb, B., & Sali, A. Comparative Protein Structure Modeling Using Modeller. Current Protocols in Bioinformatics 54, John Wiley & Sons, Inc., 5.6.1-5.6.37, 2016. doi:10.1002/cpbi.3
- wwPDB consortium. Protein Data Bank: the single global archive for 3D macromolecular structure data (2019). Nucleic Acids Research, 47, D520-D528, doi:10.1093/nar/gky949