

# Reporte técnico Docking molecular de la proteína E del SARS-CoV)

María Isabel Iñiguez Luna<sup>\*†‡</sup>

## 1 Introducción

Los estudios encaminados a los coronavirus (CoV) causantes de resfriados comunes en los seres humanos, y recientemente responsables de síndromes respiratorios agudos (SARS) (Mandala et al., 2020) han cobrado interés primario en el campo de la investigación para ampliar rápidamente los conocimientos científicos sobre el nuevo virus COVID-19, responsable actual del brote de la enfermedad por coronavirus (OMS,2021) con la intención preservar la salud y prevenir la propagación del brote se ha aprovechado los alcances tecnológicos del poder computacional para el diseño de nuevas estrategias terapéuticas para la atención de la pandemia vigente causada por el coronavirus. Las principales proteínas estructurales del coronavirus son S (pico), E (envoltura) son proteínas integrales de la membrana. Un enfoque prometedor para la prevención, son el diseño de las vacunas vivas atenuadas, y algunas de estas vacunas son dirigidas a la proteína de la envoltura (E), que es una pequeña proteína de membrana que forma canales iónicos (Surya et al., 2018). Las proteínas de la envoltura (E) de CoV son pequeñas y comprenden entre 76 y 109 aminoácidos aproximadamente, en su estructura presentan un solo domino -helicoidal transmembranal, y tienen un extremo N hidrófilo corto, y la cola hidrofóbica C-terminal, que presenta una organización pentamérica (Li et al., 2014). Los análisis in silico de acoplamiento molecular (AM), permite predecir la estructura más probable entre un ligante y el sitio de unión de una enzima específica, y así establecer una conjetura con base en la energía libre de unión de las conformaciones energéticamente más favorecidas o estables, es decir, aquellas que requieren del menor gasto energético y son más probables a ocurrir. El acoplamiento molecular consta de dos fases: la primera consiste en el docking o proceso de búsqueda de la conformación y orientación de moléculas, y la segunda etapa scoring que determinar el valor o puntaje de la interacción entre las dos partículas. Para la búsqueda de conformaciones estables se apoya en la aplicación de diversos algoritmos de búsqueda que están diseñados para predecir la actividad biológica de los compuestos seleccionados, mediante la evaluación de las interacciones entre los ligantes y los objetivos potenciales (Bartuzi et

---

\*CICE. Email: zs18024976@gmail.com

†

‡

al., 2017). El acoplamiento molecular es frecuentemente empleado en el proceso de diseño de nuevos fármacos asistido por computadora.

Descripción del problema a resolver

Identificar la secuencia que codifica a la proteína de envoltura del coronavirus SARS-CoV-2, a partir de la secuencia completa del genoma y evaluar mediante el análisis de acoplamiento molecular si la amantadina es un ligante con potencial farmacológico para la proteína E.

## 2 Metodología

Para el análisis in silico se descargó la estructura cristalográfica de la proteína E depositada en la base de libre acceso Protein Data Bank (PDB) por Torres y colaboradores, (2017) con los códigos 5x29 (NMR) y 2mm4 (NMR) depositada por Li y colaboradores (2014). Se procedió con la preparación del modelo la proteína a través del programa AutoDock 4.2.6. (GNU, General Public License) 25. En primer lugar, se eliminaron los modelos superpuestos en el archivo PDB descargado. preparación del ligante amantadina se recupero el smiles (Simplified Molecular Input Line Entry Specification) de las bases de datos Drugbank y Zinc15. Se generaron los archivos con las estructuras 2D y 3D, la adición de hidrógenos faltantes a través del software Open Babel GUI versión 2.4.1 (GNU, General Public License). El acoplamiento molecular se realizó utilizando el programa AutoDock 4.2.6 (GNU, General Public License). La configuración de los parámetros se mantuvo en los dos acoplamientos moleculares. Se utilizó el algoritmo genético Lamarckiano para la búsqueda conformacional con 2,500,000 evaluaciones por conformación de cada ligante. En total por ligante, se aplicaron 9 ejecuciones independientes, para generar las 9 conformaciones más estables. El área determinada de para el acoplamiento molecular fue 50x50x50 Å, y las coordenadas que centran al sitio activo sobre en los ejes x, y, z fueron -0.316,-0.016,-0.173 respectivamente.

## 3 Resultados

En el modelo 2mm4 se obtuvieron valores energéticos entre -4.0 y -4.4 con la formación de enlaces en los aminoácidos LEU 19, PHE 20 y PHE 23 presentado un enlace pi-sigma con la proteína de envoltura E. En el Modelo de la proteína 5x29 con cinco cadenas (A-B) valores en la tabla 2. Las conformaciones 4 y 5 se posicionaron entre las cadenas A,D ,E del modelo proteico con cinco cadenas (5x29) y los aminoácidos PHE23, LEU 19 Y PHE20. Por otro lado, las conformaciones 1,2,3,6, 7 ,8 y 9 muestran interacciones con las cadenas A,B, E del modelo proteico 5x29 entre los aminoácidos Leu 31, 28; TYR 57 y PRO54.

## 4 Conclusión

La herramienta de acoplamiento molecular nos permitió visualizar que el fármaco antiviral amantadina claramente muestra interacciones con aminoácidos presentes en las estructuras de la proteína E. Los valores energéticos obtenidos fueron muy cercanos con tan solo una 1 Kcal/mol de diferencia entre los valores medios recuperados entre la estructura de una

TABLE 1: Tabla 1. Resultados del acoplamiento molecular de la amantadina y la estructura de una cadena de la proteína E.

Fármaco	Id Zinc15	K/cal/mol
Amantadina	ZINC00000096825	-4.4
		-4.4
		-4.4
		-.2
		-4.1
		-4.1
		-4.0
		-4.0
		-4.0

TABLE 2: Tabla 2. Resultados del acoplamiento molecular de la amantadina y la estructura pentamérica de la proteína E.

Fármaco	Id Zinc15	K/cal/mol
Amantadina	ZINC00000096825	-5.4
		-5.4
		-5.3
		-5.3
		-5.3
		-5.2
		-5.2
		-5.1
		-5.1

cadena de la proteína E y la estructura pentamérica. Sin embargo, no es factible concluir que la amantadina es un buen agente farmacológico debido a que el análisis realizado se limitó a solo un fármaco antiviral. Tabla 1. Resultados de energía de unión de la amantadina y el estructura de una cadena de la proteína E.

## Referencias

- Bartuzi, D., Kaczor, A., Targowska-Duda, K., and Matosiuk, D. (2017). Recent advances and applications of molecular docking to G protein-coupled receptors. *Molecules*, 22(2), 340. Li, Y., Surya, W., Claudine, S., Torres, J. (2014). Structure of a conserved Golgi complex-targeting signal in coronavirus envelope proteins. *Journal of Biological Chemistry*, 289(18), 12535-12549. Mandala, V. S., McKay, M. J., Shcherbakov, A. A., Dregni, A. J., Kolocouris, A., Hong, M. (2020). Structure and drug binding of the SARS-CoV-2 envelope protein transmembrane domain in lipid bilayers. *Nature Structural Molecular Biology*, 27(12), 1202-1208. Surya, W., Li, Y., Torres, J. (2018). Structural model of the SARS coronavirus E channel in LMPG micelles. *Biochimica et Biophysica Acta (BBA)-Biomembranes*, 1860(6), 1309-1317.

Grice, H. P. (1975). Logic and conversation. In P. Cole & J. L. Morgan, (Eds.), *Speech Acts*, pp. 41–58. London: Academic Press. <http://dx.doi.org/3.14159--1x>.

## Appendix

The asterisk means that these divisions are not numbered.

### How to write Mathematics

This section is completely redundant with the text. Do not do that. This is just an example.

L<sup>A</sup>T<sub>E</sub>X is great at typesetting mathematics. Let  $X_1, X_2, \dots, X_n$  be a sequence of independent and identically distributed random variables with  $E[X_i] = \mu$  and  $\text{Var}[X_i] = \sigma^2 < \infty$ , and let

$$S_n = \frac{X_1 + X_2 + \dots + X_n}{n} = \frac{1}{n} \sum_i^n X_i$$

denote their mean. Then as  $n$  approaches infinity, the random variables  $\sqrt{n}(S_n - \mu)$  converge in distribution to a normal  $\mathcal{N}(0, \sigma^2)$ .

### How to create Sections and Subsections

Use section and subsections to organize your document. Simply use the section and subsection buttons in the toolbar to create them, and we'll handle all the formatting and numbering automatically.