

1. Introducción

El *Toxoplasma gondii* (Figura 1) es el parásito más común en el mundo, afectando alrededor del 30% de los seres humanos y presentando la mayor diversidad de hospederos. Su éxito radica en la habilidad para invadir cualquier célula nucleada de cualquier animal de sangre caliente; en general se encuentra en más de 350 especies de mamíferos y aves [1].

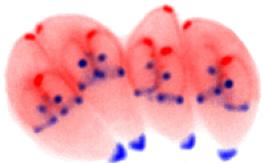


Figura 1: *Toxoplasma gondii*. Imagen recuperada de <https://doi.org/10.1371/journal.ppat.0020013>

T. gondii es el parásito causante de la infección llamada toxoplasmosis, cuyo comportamiento es todavía un misterio. Debido a que los procesos moleculares que influyen en su desarrollo son numerosos, cabe resaltar que en la actualidad se puede controlar con medicación permanente y que en individuos con sistemas inmunológicos sanos es asintomática, pero en algunas condiciones específicas esta infección puede conducir a graves desenlaces, y como no existe cura que la elimine por completo se ha convertido en una problemática de salud pública a nivel mundial [1].

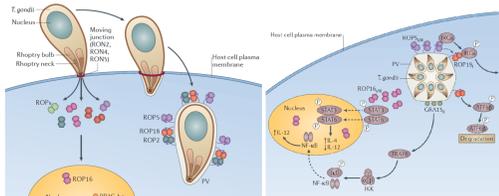


Figura 2: Modelo de infección por toxoplasmosis. Imagen recuperada de <https://www.nature.com/articles/nrmicro2858>

Una forma de entender los procesos biológicos y en particular el de la infección por toxoplasmosis, es desde una perspectiva molecular mediante el análisis de las proteínas que se involucran en el proceso estudiado (Figura 2); este análisis se puede centrar en el reconocimiento de las regiones específicas de interacción entre las proteínas del patógeno y las proteínas de célula hospedera; regiones que viabilizan el proceso de infección. Este análisis puede hacerse in silico (vía simulación computacional), basándose en la búsqueda de información en la estructura primaria de las proteínas relacionadas con la infección [2].

2. Métodos

Las técnicas experimentales para analizar las regiones de interacción en ocasiones resultan lentas, laboriosas y costosas, por ello gracias a que la estructura primaria de las proteínas están compuestas por una cadena lineal de aminoácidos, los cuales pueden ser cuantificados según sus características; es posible obtener una señal discreta que represente la proteína (Figura 3) y por ende contenga información de sus regiones de interacción. Esta representación de la proteína (señal discreta) posibilita su estudio por medio de análisis de procesamiento de señales o análisis de series temporales.

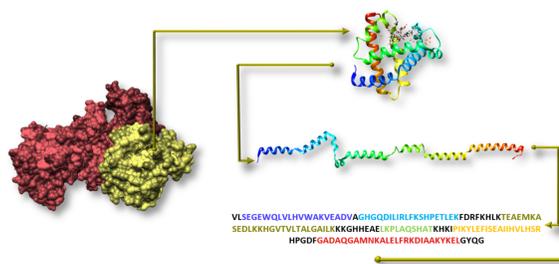


Figura 3: Estructura primaria de una proteína.

Dentro de las técnicas de procesamiento de señales, el análisis frecuencial de la señal de la proteína ha dado resultados satisfactorios como se observa en los artículos

de Wei Hu [3], Cosic [4] y Arenas [2], donde se evidencia que si dos señales de proteínas en su dominio espectral tienen picos significativos en frecuencias iguales, puede ser indicador de similitud estructural, funcional o alguna interacción (Figura 4). Sin embargo el análisis frecuencial se queda corto a la hora de reconocer las regiones específicas de interacción, si bien puede indicar que existen esas regiones, no las puede ubicar; es aquí donde el análisis tiempo-frecuencia resulta atractivo, ya que este análisis relaciona el dominio frecuencial con el dominio temporal.

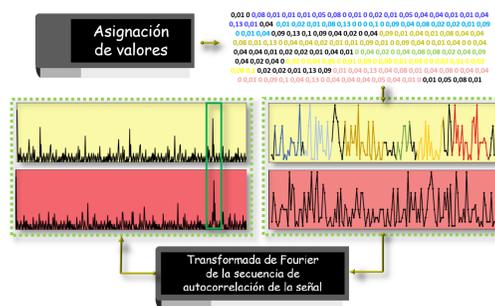


Figura 4: Señales de proteínas con sus respectivas densidades espectrales.

A través de la representación en tiempo y frecuencia $P(t, f)$ de una señal, se pueden identificar frecuencias asociadas a sitios de alta energía dentro de la señal, convirtiéndolos en posibles regiones de interacción (Figura 5).

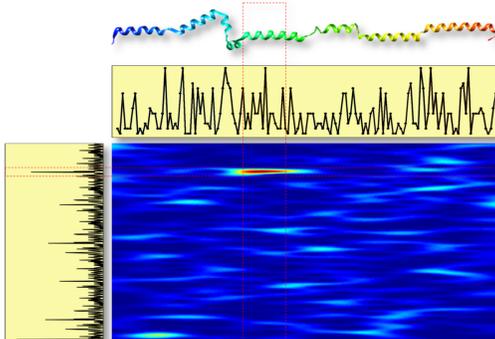


Figura 5: Espectrograma de la señal de la proteína.

Esta representación $P(t, f)$ tiene la siguiente forma [5]:

$$P(t, f) = \frac{1}{4\pi^2} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} e^{-j\theta t - j\tau f + j\theta u} \phi(\theta, \tau) x^* \left(u - \frac{1}{2}\tau \right) x \left(u + \frac{1}{2}\tau \right) dud\tau d\theta,$$

y satisface las siguientes propiedades [6]:

$$|x(t)|^2 = \int_{-\infty}^{+\infty} P(t, f) df \quad (1)$$

$$|s(f)|^2 = \int_{-\infty}^{+\infty} P(t, f) dt \quad (2)$$

$$E = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} P(t, f) df dt \quad (3)$$

donde $x(t)$ es la señal en el tiempo, $s(f)$ la densidad espectral de la señal y E la energía total de la señal. $P(t, f)$ puede llegar a revelar diferentes resultados dependiendo del kernel $\phi(\theta, \tau)$ que se utilice [7], por ello a continuación se presentan algunos:

Referencia	Kernel $\phi(\theta, \tau)$
Wigner-Ville	1
Margenau and Hill	$\cos\left(\frac{\theta\tau}{2}\right)$
Page	$e^{\frac{i\theta\tau}{2}}$
Spectrogram	$\int h^* \left(u - \frac{1}{2}\tau \right) h \left(u + \frac{1}{2}\tau \right) e^{-i\theta u} du$
Choi and Williams	$e^{-\frac{\theta^2\tau^2}{\sigma}}$
R.I.D.	$1 + \cos\left(\frac{2\pi\nu}{\tau}\right)$

3. Resultados

El "Local Variance-Switching" (Figura 6) es un programa resultado de ésta investigación, el cual gráfica distribuciones tiempo-frecuencia de cualquier señal, en especial señales de proteínas; además, puede detectar

energía localizada en un dominio dado de la distribución.

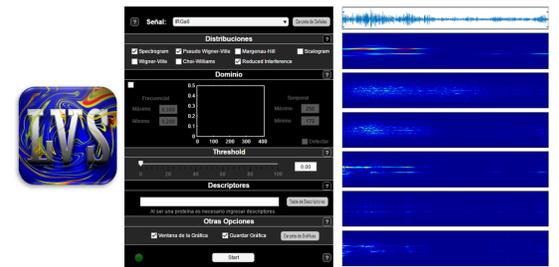


Figura 6: Local Variance-Switching.

Este programa facilitó la identificación de un péptido inhibidor de la proteína ROP5 del *T. gondii* derivado de la proteína IRGb2-b1 presente en especímenes inmunes a la Toxoplasmosis (Figura 7), el cual se validó con técnicas experimentales dando resultados satisfactorios (Figura 8).

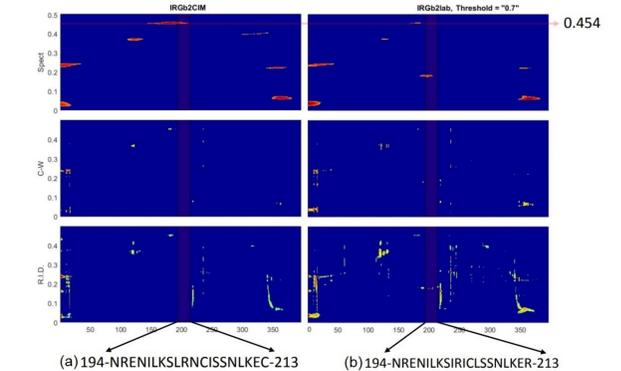


Figura 7: Espectrograma (Spect), Distribución Choi-Williams (C-W) y Distribución de Interferencia Reducida (R.I.D.) de las señales de las proteínas IRGb2CIM (inmune) y IRGb2lab (susceptible).

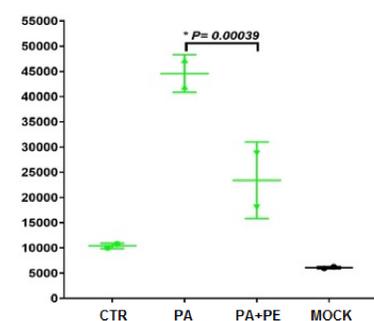


Figura 8: Ensayo de crecimiento del *Toxoplasma*. CTR: Sin parásitos, PA: Con parásitos, PE: Péptido, MOCK: PBS solución.

"Time-Frequency Approach Applied to Finding Interaction Regions in Pathogenic Proteins" [8] es el artículo producto de esta investigación, donde se expone información más detalla.

Referencias

- [1] Correa, D. (2017). Toxoplasmosis, *Revista Ciencia*, 68(1), 54-57.
- [2] Arenas, A. F.; Salcedo, G. E.; Montoya, A. M. & Gomez-Marín, J. E. (2015). MSCA: a spectral comparison algorithm between time series to identify protein-protein interactions. *BMC bioinformatics*, 16(1), 152.
- [3] Hu, W. (2010). Highly conserved domains in hemagglutinin of influenza viruses characterizing dual receptor binding. *Natural science*, 2(09), 1005.
- [4] Cosic, I.; Cosic, D.; & Lazar, K. (2017). Cancer Related BRCA-1 and BRCA-2 Mutations as Analysed by the Resonant Recognition Model. *Journal of Advances in Molecular Biology*, 1(2).
- [5] Cohen, L. (1989). Time-frequency distributions-a review. *Proceedings of the IEEE*, 77(7):941-981.
- [6] Qian S. & Chen D. (1996). Joint time-frequency analysis: methods and applications. *Prentice-Hall, Inc.*
- [7] Jeong J. & Williams W. (1992). Kernel design for reduced interference distributions. *IEEE Transactions on Signal Processing*, 40(2):402-412.
- [8] Arenas, A. F.; Arango, N.; Arenas, J. C. & Salcedo, G. E. (2019). Time-Frequency Approach Applied to Finding Interaction Regions in Pathogenic Proteins. *Bioinformatics and Biology Insights*. January. doi: 10.1177/1177932219850172.