

Conference at the faculty of Chemistry, University of Montevideo, Uruguay, 2005

## APLICACIÓN DE DIFERENTES MODELOS DE REDES NEURONALES ARTIFICIALES (RNA) A PROBLEMAS DE LA QUÍMICA MÉDICA

N.E. Campillo, A. Guerra y J.A. Páez

Instituto de Química Médica, CSIC, Juan de la Cierva 3, 28006-Madrid-España

Palabras Claves: Redes Neuronales Artificiales (RNAs), QSPR, Descriptores topológicos

Las Redes Neuronales Artificiales (RNA) son una herramienta muy útil en la solución de problemas de la Química Orgánica, especialmente en el campo de la Química Médica. Esta metodología se integra dentro de las ya clásicas relaciones cuantitativas entre estructura y actividad biológica (QSAR) o entre estructura y propiedades (QSPR), ambas muy utilizadas en Química Médica<sup>1</sup>.

El trabajo consiste fundamentalmente en el aprovechamiento de la capacidad de predicción que tienen las RNAs, predicción que se centra fundamentalmente en la actividad biológica o propiedades ADMET de nuevos fármacos. Puesto que las redes neuronales son artificios matemáticos que operan sobre números, se hace necesario convertir o codificar las estructuras a valores numéricos. Solo así podremos relacionar las estructuras con sus propiedades, que normalmente se expresan con valores numéricos.

Existen diferentes aproximaciones para describir las estructuras que se van a utilizar en un estudio QSPR. La metodología que nosotros hemos desarrollado se basa en la codificación de las estructuras desde un punto de vista topológico<sup>2</sup> empleando un programa llamado CODES<sup>®</sup> desarrollado en el IQM.

CODES genera descriptores topológicos de las estructuras mediante una red neuronal basada en un modelo de activación y competición interactiva<sup>3</sup> (IAC) a partir de las estructuras gráficas. CODES parte de la representación gráfica de la molécula, generando el correspondiente espacio topológico que posteriormente es trasladado a un espacio

neuronal (Red Neuronal Isomorfa). En la figura 1 se representa el proceso de codificación de una estructura, a través de su correspondiente Red Neuronal Isomorfa.

En la figura 1 se representa el proceso de codificación de una estructura. Los pesos de las conexiones entre las unidades ( $U_i$ ) vienen determinadas por la naturaleza del enlace entre cada par de átomos de la estructura: 1, 2 y 3, para los enlaces simple, doble y tripe respectivamente y -1 para los átomos no enlazados.

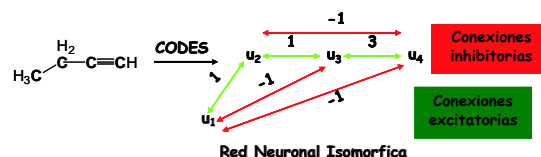


Figura 1

El color rojo y azul indica la naturaleza activante o competitiva de la conexión, respectivamente. Así, la red isomorfa en el estado inicial solo depende de la naturaleza de los átomos y de las conexiones entre los diferentes átomos de la estructura. La red neuronal isomorfa así creada se define como del tipo IAC. Su procesamiento finaliza una vez que los valores de la actividad de cada una de las fuerzas activantes y las competitivas se han equilibrado.

A lo largo de todo el proceso se van recogiendo para cada iteración los valores de activación de cada unidad o neurona, valores que constituyen la matriz dinámica (Figura 2).

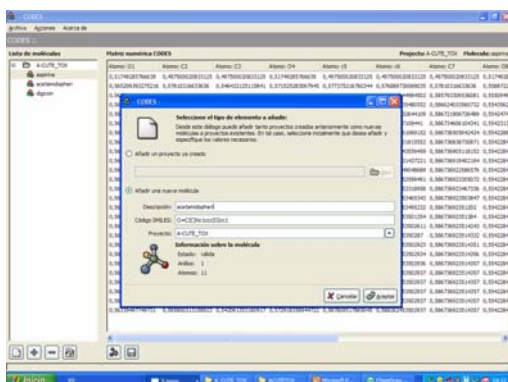


Figura 2

La figura 2 muestra a matriz dinámica de una estructura. Esta matriz es el fichero de salida del programa CODES.

La matriz esta constituida por columnas que expresan los valores de actividad de las neuronas que representan a los distintos átomos de la molécula. Cada fila refleja iteraciones consecutivas hasta llegar al equilibrio de la RNA.

En la figura 3 se esquematiza la metodología seguida para el establecimiento de QSPR mediante el uso de RNAs.

El primer paso de este proceso es generar los descriptores topológicos para cada compuesto mediante el programa CODES.

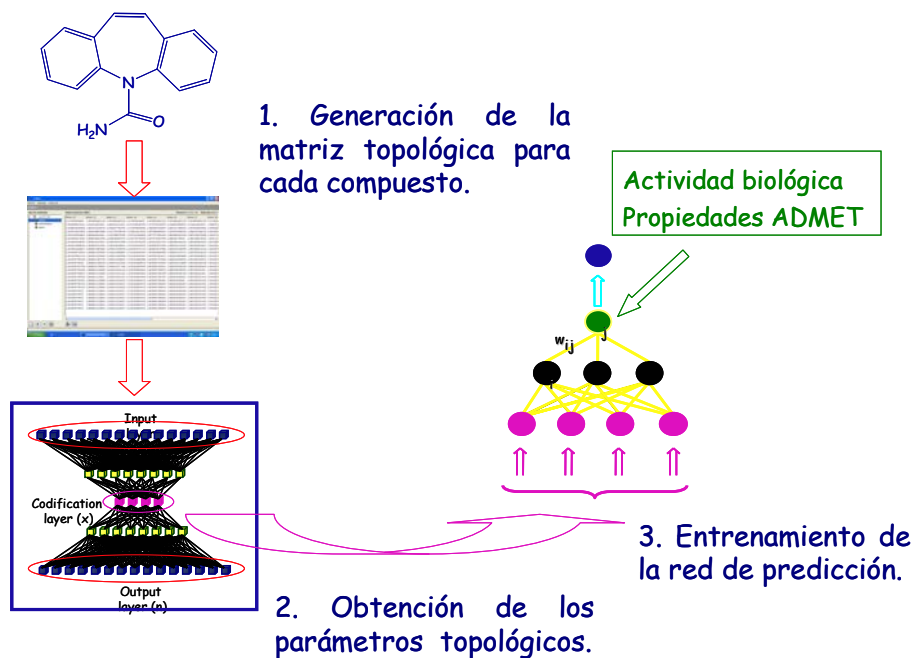


Figura 3

Las matrices dinámicas obtenidas para cada uno de los compuestos son de dimensiones muy variadas, por lo que es necesario llevar a cabo una reducción de

dimensiones para obtener el mismo número de descriptores para cada compuesto, sin que suponga pérdida de información.

La idea implícita en el concepto de reducción de dimensiones radica en disminuir la complejidad de un sistema, sin desvirtuar sus características intrínsecas ni su naturaleza. Para realizar una reducción de dimensiones utilizamos una red con una estructura simétrica  $n-z-x-z-n$  siendo  $n$  el número de átomos de cada estructura (matriz dinámica), y  $x$  el número de los nodos de la capa oculta (Figura 3). Así, estos descriptores topológicos serán los datos de entradas de modelos matemáticos generados con redes neuronales artificiales, siendo posible no solo estudiar las relaciones estructura-propiedades sino que mediante un proceso de generalización se extiende el conocimiento de estructuras conocidas a otras estructuras desconocidas.

El modelo desarrollado empleando los descriptores generados por CODES y la metodología de reducción de dimensiones proporciona una herramienta de gran utilidad en la predicción tanto de la actividad biológica como propiedades farmacocinéticas, estableciéndose por tanto como una manera sencilla y accesible de evitar costes innecesarios en el desarrollo de nuevas dianas terapéuticas<sup>4-6</sup>.

Como se ha comentado anteriormente esta metodología se puede emplear tanto para la predicción de la actividad biológica de nuevas estructuras, como para la predicción de cualquier propiedad.

Si bien en la actualidad nuestro grupo de investigación está más centrado en la predicción de propiedades ADMET tales como la absorción oral y el paso de la barrera hematoencefálica de un grupo de moléculas químicas y biodiversas.

En conclusión las ventajas más destacables de esta metodología se pueden resumir en los siguientes puntos:

- Es aplicable a conjuntos de moléculas químicas y biodiversas.
- No es necesario conocer la estructura tridimensional de los mínimos conformacionales.
- No es necesaria una selección previa de parámetros
- Es aplicable a moléculas quirales

## Agradecimientos

Queremos agradecer a Emilia Bayo, su ayuda en la realización de este trabajo.

Este trabajo está financiado por la UE mediante un proyecto integrado (FP6-512051).

## Referencias

<sup>1</sup>K.L. Peterson. In *Reviews in Computational Chemistry*; Lipkowitz, K. H. and Boyd, D. B., Ed.; Wiley-VCH: New York, vol. 16 :53-140 (2000)

<sup>2</sup>C. Ochoa, J. Rodríguez, M. Rodríguez, A. Chana, A.; M. Stud, P. Alonso-Villalobos, M. Martínez-Grueiro. *Med. Chem. Res.*, 7: 530 (1998).

<sup>3</sup>D.E. Rumelhart, J.L. McClelland, J.L. *Explorations in Parallel Distributed Processing*; MIT Press: Cambridge, MA. (1989).

<sup>4</sup>C. Gil et al. Good oral absorption prediction on non-nucleoside benzothiadiazine dioxides human cytomegalovirus inhibitors using combined chromatographic and neuronal network techniques. *Bioorg Med Chem Lett.* 1;15(7):1919-1921 (2005)

<sup>5</sup>Martínez A. et al. Prediction of Drug Half-life Values of Antihistamines Based on the CODES/Neural Network Model. *Quant.Struc.-Act.Relat.* 19:448-454 (2000)

<sup>6</sup>Ochoa, C. et al. Application of neuronal networks to the study of structure-activity relationships of 6,7-diarylpteridines as nematocides. *Med. Chem. Res.* 7, 530-545 (1998).

®CODES: idea original del Prof. Manfred Stud (CSIC).