

NUEVOS DESCRIPTORES ATOMICOS Y MOLECULARES PARA ESTUDIOS DE ESTRUCTURA-ACTIVIDAD. APLICACIONES.

M.C. Ramón Carrasco Velar, Investigador Auxiliar.

Departamento de Química, Centro de Química Farmacéutica, Calle 200 y 21,
Apartado Postal 16042, Atabey, Playa, Ciudad de La Habana, Cuba.

11 de septiembre de 2003.

Tutores: Dr. Rolando Pellón Comdom (CQF) y Dr. Luis Montero Cabrera (Facultad de Química, Universidad de la Habana).

TRABAJO PRESENTADO EN OPCION AL GRADO CIENTIFICO DE DOCTOR EN CIENCIAS QUIMICAS.

PRIMERA TESIS DE DOCTORADO DEFENDIDA EN CUBA A DISTANCIA.

Los métodos que correlacionan cuantitativamente la estructura química con la actividad biológica se engloban bajo el término QSAR (siglas de la expresión anglófona *Quantitative Structure Activity Relationships*). Uno de los intereses planteados para este trabajo fue la creación de modelos matemáticos capaces de describir cuantitativamente la actividad biológica de cefalosporinas, así como de permitir la predicción de la actividad de nuevos derivados que se pudieran sintetizar. Como antecedente de las dificultades que se le podían presentar a todo el que pretendiera modelar nuevas cefalosporinas está, entre otras, la opinión de Frere y col. con respecto a las dificultades existentes para el establecimiento de esas relaciones cuantitativas entre la estructura química y la actividad biológica de los β -lactámicos:

En cada bacteria existen diferentes PBP's (Protein Binding Penicillins) y, por tanto, diferentes constantes de velocidad de asociación.

Cada molécula de penicilina o cefalosporina tiene diferentes mecanismos de degradación en los que, a) la salida del sustituyente en R_3 depende de la naturaleza del grupo saliente; b) se encuentran diferentes velocidades de formación del complejo enzima-sustrato porque esta depende del par analizado; c) la ruptura del enlace éster en el complejo enzima-sustrato, presenta diferentes constantes de velocidad; d) hay al menos dos tipos generales de β -lactamasas, a saber: metalo-lactamasas dependientes de Zn^{2+} y serino β -lactamasas con diferente reactividad; e) el Factor de Permeabilidad, el cual se agrava en el caso de las Gram (-) por la presencia de la membrana lipídica externa (si hay β -lactamasas presentes en el periplasma, mayor influencia habrá en la difusión hasta la molécula diana); f) las amidasas responsables de inactivación de las penicilinas y cefalosporinas por ruptura del enlace amida en la posición 7, son diferentes; y por último, g) la interrelación entre estos grandes factores, los cuales dependen todos, en mayor o menor grado, de la estructura química de los compuestos.

Lo anterior significó por lo tanto, un reto para tratar de dar respuesta a esos intereses. La complejidad del problema vaticinaba una relación no lineal compleja entre la estructura y la respuesta biológica. En ese caso, el método de elección estaba constituido por las llamadas redes de neuronas (RN) en alguna de sus modalidades. Esto, sin embargo, tenía los inconvenientes propios del trabajo con redes, como lo es ante todo, el trabajar con una caja negra y desconocer el comportamiento de las variables y cómo modificarlas en el momento de pretender diseñar compuestos nuevos. Un aspecto importante a precisar con respecto a la utilización de las RN era si valía la pena perder las ventajas que brinda el análisis de regresión lineal múltiple en cuanto a simplicidad, rapidez, eficacia, facilidad de interpretación de los modelos, análisis de tendencia de las variables, etc. Sin embargo, una de las ventajas de las redes de neuronas que obliga a tenerlas en consideración era justamente la posibilidad de poder plantearse modelos teóricos no lineales complejos.

Actualmente, la forma de abordar el diseño de fármacos sigue el camino de la Química Combinatoria, pues de esa manera, se disminuyen los tiempos y costos para la obtención de un nuevo fármaco. Los métodos QSAR están en la primera parte de este enfoque como un medio para la optimización de la familia de compuestos que se desarrolla a partir de un compuesto cabeza de serie. Sin embargo, ¿son suficientes los descriptores atómicos y moleculares descritos en la literatura para enfrentar un estudio QSAR? Pudiera decirse que sí porque, según H. Wu, hay varios tipos principales de descriptores moleculares introducidos como los geométricos, los topológicos, los electrostáticos, los químico-cuánticos y los termodinámicos, pero, ¿abarcan estos descriptores todo el espectro de posibilidades de descripción de la estructura química y las propiedades a ella asociadas? Por lo tanto, se planteó como *hipótesis* de trabajo, que es posible modificar índices atómicos y moleculares de forma tal, que brinden un nuevo contenido de información acerca de las propiedades químico-físicas, para ser aplicados en estudios QSAR. A partir de esta hipótesis, se plantea el *problema científico* de ¿qué tipo de modificaciones introducir a esos índices? para lo cual, se trazó como *objetivo general*, utilizar la refractividad atómica para modificar índices topológicos conocidos, y como *objetivos específicos*: a) Transformar el Índice del Estado Electrotológico en Electrotopográfico. b) Modificar el Índice del Estado Electrotológico y el índice de Randiæ con las refractividades atómicas. c) Demostrar la aplicabilidad de estos índices en diferentes familias de compuestos y diferentes actividades biológicas, con particular énfasis en las cefalosporinas. d) Evaluar algunas capacidades de las redes de neuronas en los estudios QSAR de antibióticos β -lactámicos. Este tipo de enfoque para la modificación de índices topológicos constituye una *novedad científica* porque, a pesar de la profusión de índices atómicos y moleculares que se conocen, no se encuentra reportado el desarrollo de índices híbridos que combinen propiedades químico-físicas particionables con índices topológicos en una sola entidad. Esto se debe esencialmente a un principio establecido y aceptado por consenso en la comunidad científica, de la poca confiabilidad de los resultados por la naturaleza diversa de los términos a combinar. Los descriptores propuestos en esta tesis dan lugar a una familia de descriptores que combinan la topología de las moléculas con propiedades químico-físicas particionables como las refractividades atómicas. Al ponderar de esta manera las matrices de conectividad del grafo molecular, se logra brindar una idea de la influencia de las fuerzas de dispersión de London en fragmentos de dicha molécula, lo que sugiere su utilidad en estudios QSAR.

El Índice de Partición de la Refractividad Molecular ${}^pMR\chi$ se define por la ecuación 1 en donde la suma es sobre todos los vértices i adyacentes en el grafo, $\delta^{MR}(v_i)$ es el valor de la refractividad atómica del v_i , más los valores de la refractividad atómica de los hidrógenos enlazados al átomo v_i , para poder incluir las contribuciones al valor del índice en el grafo desprovisto de hidrógeno. El índice del Estado Refractotopológico \mathfrak{R}_i , que también se propone en este trabajo, se desarrolla a partir de la teoría del grafo químico y de la partición de la refractividad atómica definida por Ghose y Crippen. El índice se basa en la influencia de las fuerzas de dispersión de cada átomo sobre cada uno de los restantes en la molécula, modificado por la topología molecular. El R-state (\mathfrak{R}_i), para un átomo i se define por la ecuación 2, en donde AR_i es el valor de refractividad intrínseca del átomo i y ΔAR_i es un término de perturbación definida por la ecuación 3 en donde se suman todos los vértices j adyacentes en el grafo, AR_i y AR_j son los valores intrínsecos de la refractividad de los átomos i y j , respectivamente, y r_{ij}^2 es el número de átomos del camino más corto entre los átomos i y j , incluyendo tanto a i como a j .

$${}^pMR\chi = \sum [\delta^{MR}(v_i) \delta^{MR}(v_j)]^{1/2} \quad i \neq j \quad [1]$$

$$\mathfrak{R}_i = AR_i + \Delta AR_i \quad [2]$$

$$\Delta AR_i = \sum_{j=1}^N (AR_i - AR_j) / r_{ij}^2 \quad [3]$$

El Índice del Estado Electrotopográfico que aquí también se define se basa, al igual que el del Estado Electrotopológico, en el efecto electrónico de cada átomo sobre los otros átomos en la molécula. La diferencia estriba en que se considera la estructura tridimensional de la molécula. El Sstate se calcula según la ecuación 4 donde S_i es el valor del Sstate para el átomo i , I_i es un valor intrínseco asociado al átomo i y DI_i expresa el efecto perturbativo de los restantes átomos j en la molécula sobre el átomo i . El valor intrínseco I_i de cada átomo se calcula por la ecuación 5 donde N es el número cuántico principal del átomo i , d^v es el número de electrones de valencia en el esqueleto molecular ($Z^v - h$) y d es el número de electrones s en el esqueleto ($s - h$). Para cada átomo en el esqueleto molecular, Z^v es el número de electrones de valencia, s es el número de electrones en orbitales s y h es el número de hidrógenos enlazados a éste. El efecto perturbativo sobre el átomo i producido por los restantes átomos pesados presentes en la molécula se calcula según la ecuación 6 donde r_{ij} es la distancia euclídeana entre los átomos i y j tomada de la matriz de distancias, correspondiente a la configuración de mínimo energético calculada por un método semiempírico.

$$S_i = I_i + \Delta I_i \quad [4]$$

$$I_i = [(2N)^2 \delta^v + 1] / \delta \quad [5]$$

$$\Delta I_i = \sum (I_i - I_j) / r_{ij}^2 \quad [6]$$

Como consecuencia de la aplicación de estos índices, se logran establecer diferentes ecuaciones de regresión QSAR para varias actividades y familias de compuestos que mejoran los reportados, así como uno de cefalosporinas a partir de una muestra de 104 derivados. Para lograr este último, se dividió la estructura de estas últimas en tres regiones (sustituyentes en las posiciones 7 y 3 del 7-ACA y el anillo cefem). A cada uno se les calcularon los diferentes descriptores y se obtuvo un modelo con variables correspondientes a cada región de la molécula. Resalta el hecho de que, describiendo el fragmento en R7, solamente aparecen términos correspondientes al índice ${}^pMR\chi$, el cual está asociado a las fuerzas de dispersión de London. Esta zona de la molécula de cefalosporinas es la responsable de la asociación a macromoléculas en el paso de transporte y a las PBP's. Ese resultado permitió diseñar una serie de experimentación de nuevos derivados, con resultados alentadores.

Por último, se experimenta por primera vez con éxito un sistema de redes de neuronas con dos neuronas en la capa de salida correspondientes a diferentes actividades biológicas. Se trabajó con una muestra de 65 cefalosporinas con actividad reportada frente a *S. aureus* y *E. cloacae*. Para la descripción estructural de los compuestos se dividieron las moléculas en tres regiones al igual que en el análisis de regresión. A cada fragmento se les calcularon diferentes índices topológicos y topográficos entre otros descriptores. El algoritmo empleado para el establecimiento del modelo fue el de *backpropagation*. El número de neuronas en la capa oculta de la red de tres capas fue seleccionado por minimización del error relativo en el intervalo de 3 a 13 neuronas en dicha capa, por lo que la arquitectura final es 9:13:2. Se seleccionó la red con 13 neuronas en la capa oculta por ser la que mejor resultado brindó en ambas actividades. Con esa arquitectura se utilizó la técnica del *leave one out* para formar un conjunto de 65 redes cuyos valores de predicción se promediaron para dar su valor definitivo. Se obtuvieron coeficientes de correlación entre predichos y observados de 0,983 para el *S. aureus* y de 0,979 para el *E. cloacae*, utilizando los valores promedio de la predicción. Se encontraron bajos porcentajes de error en la predicción de cada bacteria los cuales son 2,6 y 2,4 % para el *S. aureus* y *E. cloacae* respectivamente. Para cada una de las neuronas de salida se encontraron valores de potencia predictiva (Q^2) de 0,966 para el *S. aureus* y de 0,958 para el *E. cloacae*. Estos elevados valores de Q^2 le confieren una elevada capacidad de predicción al modelo. Esta concepción sienta las bases para la creación de sistemas de tamizaje virtual de compuestos con actividad antibacteriana, como simulación de las baterías de tamizaje biológico experimental de antimicrobianos.

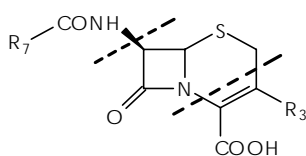


Fig. 1

El documento de tesis consta de Introducción, un capítulo de Revisión Bibliográfica y el capítulo de Resultados y Discusión en el que se exponen los resultados alcanzados en el desarrollo de los nuevos índices, su aplicación en estudios QSAR y QSPR retrospectivos. También se aborda la modelación cuantitativa y cualitativa de cefalosporinas y se desarrolla una discusión acerca de estos resultados. El documento termina con un apartado para las Conclusiones así como para las Recomendaciones. Se incluye al final la sección de Bibliografía con un total de 196 citas. Para facilitarle al lector el manejo del documento, se elaboró un Índice que contempla, además de cada epígrafe del documento, apartados para las *tablas, gráficos y figuras* por cada capítulo. Los resultados de la tesis han sido publicados parcialmente en cinco artículos en revistas nacionales e internacionales, cinco memorias de diferentes foros internacionales y expuestos en otros 33 trabajos presentados en foros nacionales o internacionales. También dio lugar a una patente cubana aprobada en el 2002.

Para el acto de defensa se realizó una conexión punto a punto entre el Centro de Química Farmacéutica en Ciudad de La Habana y la Universidad de Oriente en Santiago de Cuba, en donde se encontraban los oponentes. La velocidad de conexión fue de 31,6 Kb y se transfirieron 60 Mb de información en ambos sentidos, durante las 3,5 h que duró dicho acto. El diseño de esta modalidad para un acto de defensa de tesis doctoral estuvo a cargo del Dr. Raúl Torricella, Jefe de la Oficina de Información, CENTIC, Ministerio de Educación Superior de Cuba y en el montaje participaron entre otros, el autor de esta reseña, el Dr. B. Dupotey, J. Menéndez, R.C.Méndez, H. Alvarez y O. Rubino.