

APLICACION DE MÉTODOS DE CLASIFICACION A DI Y TRIPÉPTIDOS BIOACTIVOS CON CAPACIDAD DE INHIBICIÓN DE LA ENZIMA CONVERTIDORA DE LA ANGIOTENSINA

Andrés Pérez González, Piercosimo Tripaldi

Universidad del Azuay, Av. 24 de Mayo 7-77 y Hernán Malo. Cuenca-Ecuador

aperez@uazuay.edu.ec

Introducción

Los péptidos bioactivos poseen entre 2 y 20 aminoácidos, y actúan como compuestos moduladores y regulatorios en el organismo del ser humano, la intensidad de estas actividades dependen de la estructuración de los péptidos. (Yoshimori, et al 2010). El caso en estudio es la inhibición de la Enzima Convertidora de la Angiotensina (ECA), lo que provoca la disminución de la hipertensión, ya que al impedir que la Angiotensina I se transforme en Angiotensina II, se previene de esta forma que la bradiquina se degrade, y siga ejerciendo su efecto vasodilatador. (Chibuike & Rotimi, 2012).

Mediante la aplicación de los estudios de relación cuantitativa estructura actividad (QSAR), se puede desarrollar modelos matemáticos, a partir de resultados obtenidos de pruebas in-vitro, de la actividad de inhibición de la ECA, de péptidos específicos, y de esta forma predecir esta actividad de otros péptidos. Para el desarrollo de estos modelos se debe calcular los descriptores moleculares de las moléculas a utilizar, los cuales serán relacionados con la actividad de inhibición de la ECA

Metodología

Se construyó una base de datos de 262 moléculas de di y tripéptidos con actividad de inhibición de la ECA obtenida de la Base de Datos de Péptidos del Instituto Central de Investigación Tecnológica de Alimentos de la India (Central Food Technological Research Institute, 2014), los valores de actividad se expresan como IC_{50} .

Se asignaron los péptidos a una de tres clases determinadas, para ello se aplicó el método de agrupamiento, k-medias, los valores de los centroides obtenidos para las tres clases se muestran en la tabla 1.

Tabla 1: Centroides de las clases reales

No. Clase	Inhibición de la ECA	Valor del Centroide
1	Baja	-1.0867
2	Media	0.0948
3	Alta	0.2329

Se construyeron las moléculas utilizando el programa Hyperchem 8.0 0 (Hypercube Inc, 2008), las condiciones de trabajo se muestran en la tabla 2.

Tabla 2: Condiciones de construcción de moléculas

Aminoácidos	
Conformación	Hélice Alfa
Isómero	Levorotatorio
Campo de Fuerza Mecánico Molecular	
MM+ Componentes	Enlace
	Angulo
	Torsión
	Electroestática
	Enlaces con Hidrogenos
Optimización de la molécula	
Algoritmo	Polak-Ribiere
RMS	0.01 Kcal/(Å mol)

Los descriptores moleculares se calcularon utilizando el programa DRAGON 6.0 (Talete sr, 2010).

Para el cálculo de los modelos, se utilizará el programa Matlab R2008a, con los utilitarios “ga toolbox” y “classification_toolbox_3.1” desarrollado por el Milano Chemometrics and QSAR de la Universidad Milano Bicocca (Italia) en este utilitario se encuentran desarrollados algoritmos y métodos para desarrollo de modelos de Regresión y Clasificación.

Los métodos de clasificación utilizados fueron: KNN y CART. KNN, es un método de clasificación no paramétrico. Árbol de clasificación y regresión (CART), es un método de clasificación no paramétrico de segmentación binaria. (Deconinck et al., 2006)

Resultados

Las características del modelo generado: la tasa de error de clasificación (ET), la tasa de no error de clasificación (NET), Especificidad, Sensibilidad y Precisión, utilizando el algoritmo KNN, se muestran en la tabla 3, junto a la validación cruzada y la predicción externa. Mientras que las características del modelo generado utilizando el algoritmo CART, se muestran en la tabla 4.

Tabla 3: Modelo KNN

KNN											
Modelo Calculado				Validación Cruzada				Predicción Externa			
ET	0.09			ET	0.09			ET	0		
NET	0.91			NET	0.91			NET	1		
Clase	Especificidad	Sensibilidad	Precisión	Clase	Especificidad	Sensibilidad	Precisión	Clase	Especificidad	Sensibilidad	Precisión
1	1	0.79	1	1	0.99	0.79	0.95	1	1	1	1
2	0.93	0.98	0.92	2	0.93	0.95	0.92	2	1	1	1
3	0.98	0.97	0.97	3	0.97	0.97	0.96	3	1	1	1

Tabla 4: Modelo CART

CART											
Modelo Calculado				Validación Cruzada				Predicción Externa			
ET	0.02			ET	0.09			ET	0.01		
NET	0.98			NET	0.91			NET	0.99		
Clase	Especificidad	Sensibilidad	Precisión	Clase	Especificidad	Sensibilidad	Precisión	Clase	Especificidad	Sensibilidad	Precisión
1	0.97	1	0.86	1	0.97	0.92	0.85	1	0.99	1	0.91
2	1	0.94	1	2	0.93	0.88	0.91	2	1	0.97	1
3	0.99	1	0.99	3	0.94	0.94	0.92	3	1	1	1

El modelo generado con KNN, tiene 3 variables correspondientes a los descriptores moleculares MW, TIE y TI1_L, mientras que el modelo generado con CART, tiene 4 variables correspondientes a los descriptores moleculares MW y TIE.

El descriptor MW, es el peso molecular, en ambos modelos es la primera variable, por lo tanto podemos concluir que la actividad de inhibición de la ECA, está influenciada directamente por el tamaño del péptido, a mayor tamaño del péptido la esta actividad es mayor.

El descriptor TIE, es un descriptor parámetro topológico de estado E, que corresponde al índice de conectividad de distancia de Balaban, que se define en razón de la grado de distancia de los vértices, y no cambia en función del tamaño de la molécula o el número de anillos presentes en ella.

El descriptor TI1_L es el primer índice de MOHAR, que está relacionado con el índice de Wiener para los grafos acíclicos

Conclusiones

Luego de la aplicación de los dos métodos, se puede apreciar que las características de KNN, tienen valores más altos que los de CART, motivo por lo cual se puede considerar que el modelo KNN es más exacto y permite una aproximación más confiable, y en este caso una fácil interpretación de los descriptores relacionados por el pequeño número de estos.

CART, tiene la ventaja, de permitirnos hacer una interpretación más rápida de los resultados obtenidos, pues es más rápido de calcular, y en cada nódulo de separación de clases generado por las variables, indican claramente la condición que debe cumplirse para la asignación de clases, condición clave para realizar interpretaciones rápidas y seguras.

Considerando que los dos métodos tienen los mismos descriptores como variables, la diferencia de calidad entre ambos es mínima, por lo cual la utilización de uno u otro modelo, dependerá de los objetivos del usuario de los mismos, garantizando que va a tener buenos resultados con la utilización de cualquiera de los mismos.

Referencias

Central Food Technological Research Institute. (s.f.). Peptide Database of Central Food Technological Reserch Institute. Recuperado el 28 de Julio de 2014, de <http://www.cftri.com/pepdb/>

Chibuiké, C., & Rotimi, E. (2012). Food Protein-Derived Bioactive Peptides: Production, Processing, and Potential Health Benefits. *Journal of Food Science*, R11-R24

Deconinck, E., Zhang, M. H., Coomans, D., & Heyden, Y. V. (2006). Classification

tree models for the prediction of blood-brain barrier passage of drugs. *Journal of Quemical Information and Modeling*, 46 (3), 1410–1419.

Frank, I., & Todeschini, R. (2006). *The Data Analysis Handbook*. . Napoli-Italia .

Yoshimori, M., Eunice, L., & Jiang. (2010). *Bioactive Protein and Peptides as Funtional Foods and Nutraceutical*. Iowa-USA: Willey-Blackwell.

Hypercube Inc. (2008). *Hyperchem (Software for molecular modeling) Version 8.0*.

Talete sr. (2010). *DRAGON (Software for Molecular Descriptor Calculation)* .