

ESTUDIO QSPR INDEPENDIENTE DE LA CONFORMACIÓN PARA LA PREDICCIÓN DE CONTANTES CINÉTICAS DE OXIDACIÓN DE MICROCONTAMINANTES ACUOSOS

Pablo R. Duchowicz¹, Erlinda V. Ortiz² y Daniel O. Bennardi³

¹ Instituto de Investigaciones Físicoquímicas Teóricas y Aplicadas (INIFTA), CONICET, UNLP, Diag. 113 y 64, C.C. 16, Suc. 4, 1900 La Plata, Argentina. Correo: pabloducho@gmail.com

² IMCoDeG (CONICET), Facultad de Tecnología y Ciencias Aplicadas, Universidad Nacional de Catamarca, Maximio Victoria 55, Catamarca, Argentina.

³ Cátedra de Química Orgánica, Facultad de Ciencias Agrarias y Forestales, Universidad Nacional de La Plata (UNLP), 60 y 119, B1904AAN La Plata, Argentina

1. Introducción

En los procesos avanzados de tratamiento de agua resulta de gran interés poder identificar la presencia de microcontaminantes orgánicos (a nivel del nanogramo por mililitro), pues aún se desconocen los riesgos que ocasionan en el medioambiente o la salud humana [1]. La eficiencia de degradación de los microcontaminantes depende de su reactividad con el radical oxidrilo, y por tanto es sumamente importante conocer las constantes cinéticas de oxidación de estos compuestos [2].

A partir de información experimental disponible de la literatura [3], decidimos establecer Relaciones Cuantitativas Estructura-Propiedad (QSPR) [4] con el fin de predecir las constantes cinéticas de reacción con el radical oxidrilo (k_{OH}) en fase acuosa, asumiendo una hipotética relación entre la estructura molecular y la propiedad k_{OH} .

Los 118 microcontaminantes orgánicos poseen estructuras moleculares altamente heterogéneas, tales como fenoles, hidrocarburos aromáticos policíclicos, alcanos, compuestos aromáticos halogenados, compuestos organofosforados, etc. Con el fin de representar las estructuras químicas, se utilizan descriptores moleculares independientes de la conformación (geometría) para contemplar únicamente la naturaleza constitucional y topológica de las moléculas estudiadas.

2. Metodología

Las estructuras químicas se dibujan en formato MDL mol (V2000) en ACD/ChemSketch [5]. Luego, los programas PaDEL [6], Mold2 [7] y Epi Suite [8] permiten calcular 15251 descriptores independientes de la geometría. Estos programas son gratuitos y de libre acceso.

Se divide el conjunto molecular en un conjunto de calibración (cal) con $N_{cal} = 89$ moléculas (75%) para seleccionar los descriptores y parámetros ajustables del modelo de regresión lineal multivariable (MLR), y otro de validación (val) con $N_{val} = 29$ moléculas para determinar la capacidad predictiva del modelo obtenido. Para este fin, se utiliza la técnica "Método de Subconjuntos Balanceados" (BSM) [9].

Los modelos QSPR se establecen con un algoritmo matemático muy útil basado en la técnica MLR, el "Método del Reemplazo" (RM) [10], que permite seleccionar los mejores descriptores a partir un gran número de variables, de manera que se minimice la raíz cuadrada de la desviación cuadrática media de calibración (RMS_{cal}).

Aparte de utilizar un conjunto de validación externo, cada modelo se valida por medio de validación cruzada 'Dejar-Uno-Afuera' (*loo*) y Aleatorización-Y (*rand*) [11].

Todos los algoritmos de Matlab [12] utilizados se desarrollaron en nuestro grupo de investigación y están disponibles.

3. Resultados

La mejor relación cuantitativa estructura-propiedad encontrada a través de la técnica RM involucra a los siguientes 5 descriptores moleculares:

$$\log k_{\text{OH}} = -0.43 \text{AATS0e} + 1.50 \text{SIC5} - 0.91 \text{PC200} - 0.33 \text{KC3738} - 0.03 \text{AC.CX7} + 11.91$$

$$N_{\text{cal}} = 89, R_{\text{cal}}^2 = 0.88, \text{RMS}_{\text{cal}} = 0.21$$

$$R_{\text{ijmax}}^2 = 0.11, R_{\text{loo}}^2 = 0.85, \text{RMS}_{\text{loo}} = 0.24, \text{RMS}^{\text{rand}} = 0.45$$

$$N_{\text{val}} = 29, R_{\text{val}}^2 = 0.87, \text{RMS}_{\text{val}} = 0.11$$

Aquí, R es el coeficiente de correlación y R_{ijmax} la correlación máxima entre pares de descriptores. La capacidad predictiva de la ecuación QSPR es satisfactoria según sugieren los parámetros estadísticos de calibración, validación cruzada y externa ($R \geq 0.7$) y Aleatorización-Y ($\text{RMS}^{\text{rand}} > \text{RMS}_{\text{cal}}$).

El modelo presentado mejora al publicado por Jin y col. [3], el cual posee 7 descriptores algunos de ellos dependientes de la geometría molecular, y que se obtiene luego de eliminar a 2 moléculas con malas predicciones: $R_{\text{cal}}^2 = 0.86$, $\text{RMS}_{\text{cal}} = 0.18$, $R_{\text{val}}^2 = 0.87$, $\text{RMS}_{\text{val}} = 0.26$.

Los descriptores independientes de la conformación que posee nuestro modelo QSPR pueden calcularse fácilmente. Estas variables no presentan multicolinealidad, evidenciado por la ausencia de correlación entre un descriptor dado y los restantes de la ecuación lineal.

Finalmente, se demuestra que las 29 moléculas de validación están contenidas dentro del dominio de aplicabilidad del modelo QSPR hallado, por lo que sus predicciones pueden considerarse confiables.

4. Conclusiones

El modelo establecido en este trabajo tiende a predecir acertadamente las constantes cinéticas de reacción con el radical oxidrilo de los microcontaminantes acuosos. La importancia del estudio actual radica en la habilidad de establecer un modelo QSPR que dependa solamente de los aspectos constitucionales y topológicos de las estructuras moleculares y no de sus aspectos geométricos. Además, no plantea la eliminación de moléculas, e incluye solamente 5 descriptores moleculares. Mediante el estudio QSPR desarrollado, es posible predecir la susceptibilidad de degradación de los microcontaminantes en procesos de tratamiento de agua.

Agradecimientos. Se agradece a CONICET (PIP0151).

Referencias

- [1] Luo, Y., Guo, W., Ngo, H. H., Nghiem, L. D., Hai, F. I., Zhang, J., Liang, S., and Wang, X. C., 2014. *Sci. Total Environ.* 473-474, 619-641.
- [2] Bagheri, M. and Mohseni, M., 2015. *J. Hazard. Mater.* 294, 1-8.
- [3] Jin, X., Peldszus, S., and Huck, P. M., 2015. *Chemosphere* 138, 1-9.
- [4] Hansch, C. and Leo, A., 1995. *Exploring QSAR. Fundamentals and Applications in Chemistry and Biology.* American Chemical Society. Washington, D. C.
- [5] ACD/ChemSketch, 2016. www.acdlabs.com

XXXI Congreso Argentino de Química

25 al 28 de Octubre de 2016 Asociación Química Argentina

Sánchez de Bustamante 1749 – Ciudad de Buenos Aires – Argentina

The Journal of The Argentine Chemical Society Vol. 103 (1-2) January – December 2016 ISSN: 1852 -1207

Anales de la Asociación Química Argentina AAQAE 095 - 196

- [6] PaDEL, 2016. <http://www.yapcwsoft.com>
- [7] Hong, H., Xie, Q., Ge, W., Qian, F., Fang, H., Shi, L., Su, Z., Perkins, R., and Tong, W., 2008. J. Chem. Inf. Model. 48, 1337-1344.
- [8] Epi Suite 4.11, 2016. U.S. EPA, <https://www.epa.gov/tsca-screening-tools/epi-suite-tm-estimation-program-interface>
- [9] Rojas, C., Duchowicz, P. R., Tripaldi, P., and Pis Diez, R., 2015. J. Chromatogr. A 1422, 277-288.
- [10] Duchowicz, P. R., Castro, E. A., and Fernández, F. M., 2006. MATCH Commun. Math. Comput. Chem. 55, 179-192.
- [11] Gramatica, P., 2007. QSAR Comb. Sci. 26, 694-701.
- [12] Matlab 7.0., The MathWorks, 2016. Masachussetts, USA, <http://www.mathworks.com>