Statistical Approach of Structure-Activity Relationships: A Case Study

Sorana D. BOLBOACĂ[1], Lorentz JÄNTSCHI[2]

[1]"Iuliu Hațieganu" University of Medicine and Pharmacy Cluj-Napoca, 6 Louis Pasteur, 400349 Cluj-Napoca, Cluj, Romania. E-mail: sbolboaca@umfcluj.ro

[2]Technical University of Cluj Napoca, 103-105 Muncii Bvd., 400641 Cluj-Napoca, Cluj, Romania. E-mail: lori@academicdirect.org

Abstract

An integrated system called Molecular Descriptors Family on Structure-Activity Relationships (MDF-SAR) was developed and used for analysis and quantification of the link between compounds' structure and measured activity/property in order to obtain structure-activity relationships (SARs). The MDF-SAR approach is able to obtain simple as well as multiple linear regression models between structure (from quantum based descriptors [1] and measured activity/property using a genetic algorithm. The results obtained by using the MDF-SAR approach are online available [2]. A series of methods were proposed for assessment of the obtained models [3]. Starting with the proposed methods some client-server statistical software applications were developed. The statistical approach of structure-activity/property relationships included correlation analysis (Pearson, Spearman, Kendall and Gamma coefficients as parameters and associated significance levels), regression analysis (leave-one-out cross-validation and determination coefficients), and other inferential statistics (cross correlation coefficients, training vs. test experiment, correlated correlations analysis).

Keywords:

Structure - activity relationships; inferential statistics; models assessment

References:

[1] Jäntschi L, Bolboacă SD. Int J Quant Chem, 107(8), 1736-1744, **2007**.

[2] Jäntschi L, Bolboacă SD, Diudea MV. Int J Mol Sci, 8(11), 1125-1157, **2007**.

[3] Bolboacă SD, Jäntschi L, Env Chem Lett, DOI 10.1007/s10311-007-0119-9.