New advanced tool for the computational estimation of mutagenesis

S. Gómez and R. Gozalbes*

ProtoQSAR SL (www.protoqsar.com), Centro Europeo de Empresas Innovadoras, Parque Tecnológico de Valencia, Spain
E-mail: rgozalbes@protoqsar.com

The assessment of mutagenicity is an essential component for the evaluation of the toxicological profile of chemicals. Mutagenic effects caused by chemical agents are usually detected by the Ames test, which nowadays can be considered as a paradigm for the development of current *in vitro* toxicology.

Quantitative structure-activity relationships (QSAR) models constitute a rational and well-established methodology for prediction of chemical properties with computers. QSARs involve the construction of mathematical models relating –by means of statistical tools- the chemical structure of a series of molecules with a physicochemical/biological property. Once a correlation is established, it can be used to predict the behavior of new molecules. The use of QSARs with regulatory purposes is recognized by international agencies and institutions such as ECHA or the OECD, and for example the evaluation of pharmaceutical impurities is commonly done by companies and consultants following the requirements of the International Conference on Harmonisation (ICH) guides.

Several pieces of software are commercially available for the computational ("*in silico*") evaluation of mutagenicity. The accuracy of Ames mutagenicity prediction by these programs is typically 70-75%, which is not far from the inter-laboratory reproducibility of the Ames test (usually estimated at 80 – 85%).These tools have different drawbacks: most of them are "black boxes" for users, in which there is no degree of certainty or reliability of their predictions, it is difficult for non-specialized personnel to understand their functioning, and only experienced staff can work with it.

We will present here a completely new QSAR model for mutagenicity estimation, based on the hugest dataset constructed until now from reliable data. This model presents outstanding advantages over the current programs, such as good statistic performances, higher applicability domain and ease of use.