LQTA-QSAR: A NEW 4D-QSAR METHODOLOGY

*João Paulo A. Martins, Euzébio G. Barbosa, Kerly F. M. Pasqualoto, Márcia M. C. Ferreira**

Laboratory for Theoretical and Applied Chemometrics, Institute of Chemistry,
University of Campinas, Campinas, SP 13084-971, Brazil,
*e-mail: marcia@iqm.unicamp.br

The new 4D-QSAR approach presented and named LQTA-QSAR *(Laboratório de Quimiometria Teórica e Aplicada)*, is based on the generation of a conformational ensemble profile, CEP, for each compound, followed by the calculation of 3D descriptors. This new methodology explores jointly the main features of CoMFA and 4D-QSAR paradigms. GROMACS free package is used for molecular dynamics, MD, simulations and generating CEP. The module LQTAgrid calculates intermolecular interaction energies at each grid point considering different probes and all aligned conformations from MD simulations. These interaction energies are the descriptors employed in the QSAR analysis. The ordered predictor selection, OPS, algorithm2 recently developed in our laboratory, is applied as the variable selection method in the construction of the PLS models. OPS method has been proved to be fast and capable of providing suitable variables for the QSAR analysis. LQTA-QSAR models are thoroughly validated applying the leave-*N*-out cross-validation and y-randomization methods. The comparison of the proposed methodology to other 4D-QSAR and CoMFA formalisms was performed using a set of forty-seven glycogen phosphorylase b inhibitors (data set 1) and a set of forty-four MAP p38 kinase inhibitors (data set 2). The QSAR models were built using the OPS algorithm for variable selection. Model validation was carried out applying y-randomization and leave-*N*-out cross-validation in addition to the external validation. PLS models for data sets 1 and 2 provided the following statistics: $q^2 = 0.72$, $r^2 = 0.81$ for 12 variables selected and 2 latent variables; and, $q^2 = 0.82$, $r^2 = 0.90$ for 10 variables selected and 4 latent variables, respectively. Visualization of the descriptors in 3D space was successfully interpreted from the chemical point of view, supporting the applicability of this new approach in rational drug design.
LQTA-QSAR is available at http://lqta.iqm.unicamp.br

References

[1] Martins JP; Barbosa E; Pasqualoto KF; Ferreira MMC, LQTA-QSAR: a new 4D-QSAR methodology. *J. Chem. Inf. Comput.* Mod. 2009, in press.

[2] Teófilo RF; Martins JP; Ferreira MMC, Sorting variables by using informative vectors as a strategy for feature selection in multivariate regression. *J. Chemometr.*, 2009, 23, 32.

Acknowledgements: FAPESP, CNPq, CAPES