Molecular Informatics 2016 vol.35 N2, pages 70-80

Predictive Models for Halogen-bond Basicity of Binding Sites of Polyfunctional Molecules

Glavatskikh M., Madzhidov T., Solov'ev V., Marcou G., Horvath D., Graton J., Le Questel J., Varnek A.
Kazan Federal University, 420008, Kremlevskaya 18, Kazan, Russia

Abstract

© 2016 Wiley-VCH Verlag GmbH & Co. KGaA.Halogen bonding (XB) strength assesses the ability of an electron-enriched group to be involved in complexes with polarizable electrophilic halogenated or diatomic halogen molecules. Here, we report QSPR models of XB of particular relevance for an efficient screening of large sets of compounds. The basicity is described by pKBI2, the decimal logarithm of the experimental 1 : 1 (B :I2) complexation constant K of organic compounds (B) with diiodine (I2) as a reference halogen-bond donor in alkanes at 298K. Modeling involved ISIDA fragment descriptors, using SVM and MLR methods on a set of 598 organic compounds. Developed models were then challenged to make predictions for an external test set of 11 polyfunctional compounds for which unambiguous assignment of the measured effective complexation constant to specific groups out of the putative acceptor sites is not granted. At this stage, developed models were used to predict pKBI2 of all putative acceptor sites, followed by an estimation of the predicted effective complexation constant using the ChemEqui program. The best consensus models perform well both in cross-validation (root mean squared error RMSE=0.39-0.47logKBI2 units) and external predictions (RMSE=0.49). The SVM models are implemented on our website (http://infochim.u-strasbg.fr/webserv/VSEngine.html) together with the estimation of their applicability domain and an automatic detection of potential halogen-bond acceptor atoms.

http://dx.doi.org/10.1002/minf.201500116

Keywords

applicability domain, ensemble modeling, Halogen bonding, QSPR