First International School-Seminar

«From empirical to predictive chemistry»

Program
Book of Abstracts

Kazan
November 28-29, 2014

210 years
KFU

Vitaly P. Solov'ev[1]
A. Varnek[2]

Modelling of hydrogen bond thermodynamics

[1] Institute of Physical Chemistry and Electrochemistry, Russian Academy of Sciences, Leninskiy prospect, 31/4, 119071 Moscow, Russian Federation
[2] Laboratoire d'Infochimie, UMR 7177 CNRS, Université de Strasbourg, 4, rue B. Pascal, Strasbourg, 67000, France
solovev-vp@mail.ru

The hydrogen bond (H-bond) is one of the fundamental interactions between molecules and is of significant importance for many properties, as well as for processes of livingand abiotic nature.

In the case of intermolecular interactions, a hydrogen bond is formed between the molecule containing the X-H, referred as the H-bond donor (HBD), and the molecule containing the Y atom with which the X-H forms a bond, referred as the H-bond acceptor (HBA). As a rule, the X and Y atoms are O, N, S, Se or F.

The thermodynamic quantities (the stability constant $\log K$, the Gibbs free energy ΔG and the enthalpy ΔH) of the 1:1 (HBD:HBA) H-bond complexation in CCl_4 at 298 K are widely applicable for a quantitative assessment of H-bond strength.

Earlier, the modeling of thermodynamic quantities of the 1:1 H-bond complexes has already been attempted through various approaches such as quantum chemical methods, linear free-energy relationships (LFERs), empirical correlations and quantitative structure-property relationships (QSPR) [1,2]. The LFERs models consider $\log K$, ΔG and ΔH of the H-bond complexation as a product of the acceptor and donor parameters. In the large, used measured $\log K$, ΔG and ΔH values were not experimentally attributed to H-bond sites involved in the complexation. Thereby, the parameters encompasse the whole molecule. Then authors introduced second level of empirical descriptions for assigning the parameters to H-bond sites.

Recently, QSPR modeling of $\log K$ for the 1:1 complexation of 537 organic HBAs with 4-fluoro-phenol as the reference HBD was peformed applying new ISIDA fragment descriptors able to describe "local" properties related to selected atoms or functional chemical groups [2]. The $\log K$ values for the modeling were measured by Fourier transform IR spectroscopy (FTIR) in CCl 4 at 298 K and experimentally attributed to HBA sites taking part in the complexation. Unlike previous QSPR studies of H-bond complexation, the models based on the FTIR data and new ISIDA descriptors are able to predict the H-bond strength (basicity) of the O, N and S atoms in different hybridization states as HBAs of the polyfunctional molecules. QSPR models were obtained using support vector machine (SVM) and ensemble multiple linear regression (ISIDA MLR) methods. The models were validated with cross-validation procedures and with two external test sets. The root mean squared error of the 5-fold cross validation ISIDA MLR predictions is 0.24 of $\log K$ units. For predictions of the HBA strength of organic molecules, the SVM consensus model is freely available on web server: http://infochim.u-strasbg.fr/webserv/VSEngine.html.

1. Varnek A., Fourches D., Hoonakker F., Solov'ev V. *J. Comp.-Aided Mol. Design*, **2005**, 19, 693-703.
2. Ruggiu F., Solov'ev V., Marcou G., Horvath D., Graton J., Le Questel J.-Y., Varnek A. *Mol. Inf.*, **2014**, 33, 477–487.