

Book of Abstracts
and Scientific Program



International School-Seminar on Computer-Aided Molecular Design

May 18-20, 2016
Kazan, Russia

V. Solov'ev¹
M. Glavatskikh^{2,3}
D. Horvath²
G. Marcou²
A. Varnek²

THERMODYNAMICS OF HYDROGEN BONDING: FROM EXPERIMENTAL MEASUREMENTS TO STRUCTURE-PROPERTY MODELING

¹ A.N.Frumkin Institute of Physical Chemistry and Electrochemistry,
Russian Academy of Sciences, Leninskiy prosp., 31, 119071
Moscow, Russia;

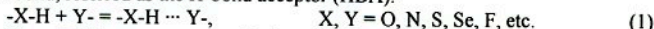
² Laboratoire de Chimoinformatique, UMR 7140 CNRS, Université
de Strasbourg, 1, rue Blaise Pascal, 67000 Strasbourg, France;

³ Laboratory of Chemoinformatics and Molecular Modeling, Butlerov
Institut of Chemistry, Kazan Federal University, Kremlevskaya 18,
Kazan, Russia;

solovev-yp@mail.ru

Hydrogen bonding is of crucial importance for many properties, as well as for processes of living and abiotic nature. For instance, it is responsible for structure formation phenomena in biological systems and artificial self-assembling systems. A quantitative assessment of H-bond strength has for a long time been important for the chemical community [1].

A hydrogen bond (H-bond) is formed between the molecule containing the X-H atoms, referred as the H-bond donor (HBD), and the molecule containing the Y atom with which the H-X forms a bond, referred as the H-bond acceptor (HBA):



The stability constant $\log K$, the Gibbs energy ΔG and the enthalpy ΔH of the 1:1 (HBD:HBA) H-bond complexation are widely used for a quantitative assessment of H-bond strength.

In this study, we discuss results of experimental measurements of H-bond with using IR and UV spectrophotometry, NMR spectroscopy and calorimetry for calculations of thermodynamic quantities $\log K$, ΔG and ΔH of hydrogen bonding in organic solution. We conclude, that known data level makes it possible to extensive QSPR (Quantitative Structure-Property Relationships) modeling allowing very quick and precise calculation of the strength or thermodynamic quantities of hydrogen bonding, that would be applicable for the wide variety of the H-bond donors and acceptors.

The estimation of the H-bond strength is based on the Gibbs energy ΔG for the 1:1 hydrogen bond complexation (1) of different HBDs with different HBDs, which represent for up to now the largest, structurally diversified dataset at 298 K in CCl_4 . The Gibbs energy for the dataset of 3388 reactions (1) was modeled by Support Vector Machine and Multiple Linear Regression methods using ISIDA Fragment descriptors. Successfully cross-validating models were then challenged to make predictions for an external test set of 641 H-bond complexes. The best consensus models perform well both in external cross-validation (root mean squared error $RMSE = 1.55 - 2.35$ kJ/mol) and external test set predictions with the $RMSE$ within the range 2.17 - 5.56 kJ/mol depending on the class of Applicability Domain.

A predictor of ΔG for hydrogen bond complexation of various organic donors with various acceptors was created on the base of the entire training set and the best models. The SVM consensus model is publically available on the server: <http://infochim.u-strasbg.fr/webserv/VSEngine.html>.

I. Ruggiu F., Solov'ev V., Marcou G., Horvath D., Graton J., Le Questel J.-Y., Varnek A. *Mol. Inf.* 2014, 33, 477-487.

VS thanks Organizers of International School-Seminar on Computer-Aided Molecular Design for support.
