

International Conference on Cheminformatics and Computational Chemical Biology July 11-12,  
2016 Brisbane, Australia



**Vitaly Solovev**

Russian Academy of Sciences, Russia

Title: Metal ions binding in water: Tools for organic ligands design

## Biography

## Abstract

Several tools for design of metal ion binders in water were realized to apply the consensus QSPR models based on sub-structural molecular fragments (SMF) as descriptors: property predictor, generator of virtual combinatorial libraries and interactive designer of compounds. The developed consensus models (CM) for predicting stability constants (log K) of the metal ion – organic ligand complexation are integrated in Forecast by Molecular Fragments (FMF) predictor. Ligands can be submitted as an SD file. The predicted log K are evaluated as an arithmetic mean of values obtained by numerous individual Multiple Linear Regression models excluding those leading to outlying values and being outside applicability domain (AD) of individual models. Three types of AD definitions can be used simultaneously or separately: bounding box, fragment control and “quorum control”. Outlying predictions of some individual models are excluded from the CM by Thompson’s rule. Chemical editor Ed-ChemS includes a generator of virtual combinatorial libraries named Combi-Lib. various libraries are generated by attaching substituent to molecular scaffolds. Then the log K values for generated compounds are estimated by the FMF predictor. An interactive designer of organic compounds is realized by interaction of the chemical editor EdChemS with the FMF predictor using coloring of atoms of chemical formula according the SMF contributions. If molecular structure is edited on the screen by EdChemS, the FMF predicts the property interactively using loaded CM. Atoms of molecular formula are colored according to the SMF contributions of CM. Color depth of atom is double sum, where first sum includes contributions of molecular fragments containing given atom, and second sum includes all individual models of CM. The fragments and their contributions are convenient tools for the rationale design of the ligands with desirable thermodynamic stability of their complexes: the data manager EdiSDF estimates mean-fragment contributions according to a set of individual models in CM. The tools use the developed QSPR models for the stability constants log-K of the 1:1 (M: L) complexes of metal ions (M) with different classes of organic ligands (L) in aqueous solution at 298 K and an ionic strength 0.1 M. The CM were prepared by the ISIDA/QSPR program for 42 metal ions: Li<sup>+</sup>, Na<sup>+</sup>, K<sup>+</sup>, Be<sup>2+</sup>, Mg<sup>2+</sup>, Ca<sup>2+</sup>, Sr<sup>2+</sup>, Ba<sup>2+</sup>, Al<sup>3+</sup>, Ga<sup>3+</sup>, In<sup>3+</sup>, Pb<sup>2+</sup>, Y<sup>3+</sup>, La<sup>3+</sup>, VO<sup>2+</sup>, Mn<sup>2+</sup>, Fe<sup>2+</sup>, Fe<sup>3+</sup>, Co<sup>2+</sup>, Ni<sup>2+</sup>, Cu<sup>2+</sup>, Ag<sup>+</sup>, Zn<sup>2+</sup>, Cd<sup>2+</sup>, Hg<sup>2+</sup>, Ce<sup>3+</sup>, Pr<sup>3+</sup>, Nd<sup>3+</sup>, Sm<sup>3+</sup>, Eu<sup>3+</sup>, Gd<sup>3+</sup>, Tb<sup>3+</sup>, Dy<sup>3+</sup>, Ho<sup>3+</sup>, Er<sup>3+</sup>, Tm<sup>3+</sup>, Yb<sup>3+</sup>, Lu<sup>3+</sup>, Th<sup>4+</sup>, UO<sub>2</sub><sup>2+</sup>, NpO<sub>2</sub><sup>+</sup> and Am<sup>3+</sup>. Studied ligands are molecules of various organic classes and data sets from 883 (Cu<sup>2+</sup>) to 28 (Am<sup>3+</sup>) organic ligands. The models have reasonable prediction performance: root-mean squared error in external 5-fold cross-validations varies from 0.49 (Li<sup>+</sup>) to 2.30 (In<sup>3+</sup>) (the log K units) which is close to observed experimental systematic errors.