# XVIII МЕНДЕЛЕЕВСКИЙ СЪЕЗД ПО ОБЩЕЙ И ПРИКЛАДНОЙ ХИМИИ

Москва, 23–28 сентября 2007 г.

## ТЕЗИСЫ ДОКЛАДОВ

В пяти томах

*Том 5*

*IV Российско-французский симпозиум «Супрамолекулярные системы в химии и биологии»*

*II Российско-индийский симпозиум по органической химии*

*Международный симпозиум по современной радиохимии «Радиохимия: достижения и перспективы»*

*Международный симпозиум «Зеленая химия, устойчивое развитие и социальная ответственность химиков»*

*Симпозиум «Нуклеофильное замещение водорода в ароматических системах и родственные реакции»*

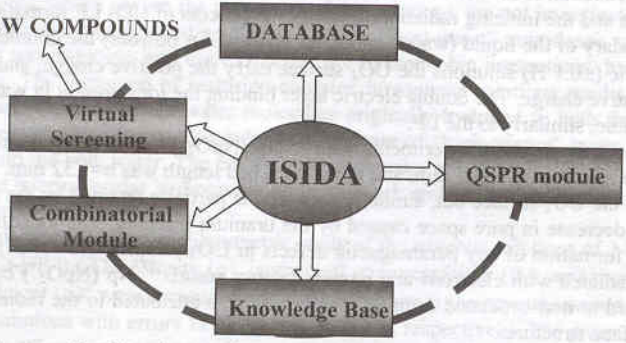
Москва – 2007

# COMPLEXATION OF METALS WITH ORGANIC LIGANDS IN WATER: COMPUTER AIDED DESIGN OF NEW METAL BINDERS

**Kireeva N.,[ab] Marcou G.,[a] Soloviev V.P.,[b] Fourches D.,[a] Varnek A.[a]**

[a]*Universite Louis Pasteur, 4, rue B.Pascal, Strasbourg, France*
[b]*Institute of Physical Chemistry and Electrochemistry, RAS, Leninsky pr-t, 31, Moscow, Russia*

Chemoinformatics technologies can be efficiently used for establishing quantitative relationships between structure and metal binding affinity of organic molecules. For this purpose we used the ISIDA (In Silico Design and Data Analysis) program package, which includes three main elements: *(i)* a database manager, *(ii)* a QSPR module for quantitative structure-property modelling and *(iii)* a generator of virtual combinatorial libraries.




Several popular machine learning methods – Multiple Linear Regression Analysis, *k* Nearest Neighbours, Support Vector Machines, Associative Neural Netwotks – implemented in the ISIDA package have been used to assess stability of the complexes of metal cations (M) with some organic ligands (L) in water using substructural molecular fragments as descriptors. Several original variable selection methods were implemented.

The models linking stability constants (log*K*) and molecular structure have been obtained for data sets of 130, 122, 235 and 1893 molecules forming M:L = 1:1 complexes with Sr, Ca, Ba and Cu(II) cations, respectively. The experimental log*K* values were critically selected from the IUPAC database and scaled to "standard" conditions (T = 25°C, I = 0.1 M). These structurally diverse data sets include derivatives of monocarboxylic, phosphoric and aminophosphonic acids, mono- and dipodands, crown-ethers and aza-crown ethers, cryptands and some others.

Predictive ability of the models was analyzed using external 5-fold cross validation procedure, in which every compound in the initial data set takes part of one of five test sets.

Developed here structure – log*K* models can be used for *"in silico"* design of new selective ligands for metal cations.