

# IMPLEMENTATION OF ORBITAL GRAPH MODEL FOR PARTIAL ATOM CHARGES DEFINITION IN VIRTUAL SCREENING TECHNIQUE

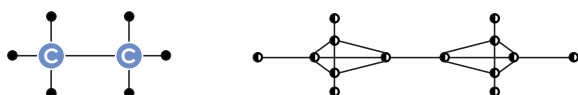
Olexander Ya. Yakovenko, Andriy G. Golub, Volodymyr G. Bdzholo and Sergiy M. Yarmoluk

Institute of Molecular Biology and Genetics of NAS of Ukraine  
Department of Combinatorial Chemistry of Biologically Active Compounds  
03143, Ukraine, Kyiv, Zabolotnogo str. 150  
E-mail: sergiy@yarmoluk.org.ua Web-site: http://www.yarmoluk.org.ua

## ABSTRACT

Partial atom charges are one of the key features of ligand description in molecular docking approach of virtual screening technique. We implemented partial atomic charge definition model originally proposed by Oliferenko A. A. [1] into virtual screening engine. This model represents atom orbital network via electric topology graph and calculates atom partial charges by equalization of atom's orbital electronegativity. The accuracy of molecular docking method was compared for charges derived from orbital graph model, molecular graph model and AM1 semi empirical approach.

## THEORY



Ohm's law:

$$\sum_{i=1}^n Y_{ij} \Phi_i = J_j$$

where  $Y$  is the matrix of nodal conductivity,  $\Phi$  is the column matrix of the potentials and  $J$  is the column matrix of the nodal currents.

The solution in matrix form is:

$$\begin{pmatrix} \Phi & Y^{-1} J \\ S & D & A & I \end{pmatrix}$$

Where  $D$  is the diagonal matrix of graph vertex,  $A$  is the adjacency matrix of the network graph and  $I$  is the identity matrix of the same order as  $A$  and  $D$ .

$$X = S^{-1} X^0$$

where  $X$  is the equalized electronegativity and  $X^0$  is the table electronegativity.

Orbital graph model:

$$q_i = \frac{E}{n} - \frac{IP}{2} - \frac{EA}{2} - \frac{E}{n^2} - \frac{IP}{2} - \frac{EA}{2} - \frac{E}{n^2} - \frac{IP}{2} - \frac{EA}{2} - \frac{E}{n^2} - \frac{IP}{2} - \frac{EA}{2}$$

Where  $D$  is the normalization factor,  $q_i$  is the orbital charge,  $E$  is the orbital electronegativity,  $IP$  is the orbital hardness,  $E$  is the atomic energy,  $IP$  is the ionization potential and  $EA$  is the electron affinity.

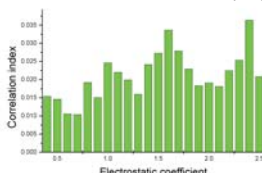
Molecular graph model:

$$q_i = \frac{E}{m} - \frac{IP}{2} - \frac{EA}{2} - \frac{E}{m^2} - \frac{IP}{2} - \frac{EA}{2} - \frac{E}{m^2} - \frac{IP}{2} - \frac{EA}{2} - \frac{E}{m^2} - \frac{IP}{2} - \frac{EA}{2}$$

where  $m$  is the normalization factor,  $q_i$  is the atom charge and  $X^0$  is the atom table electronegativity

## RESULTS

The highest accuracy was obtained for molecular graph charge definition model. The accuracy values were calculated as correlation index of docking score with natural logarithm of half of the rest ferment activity in presence of 30 M inhibitor.

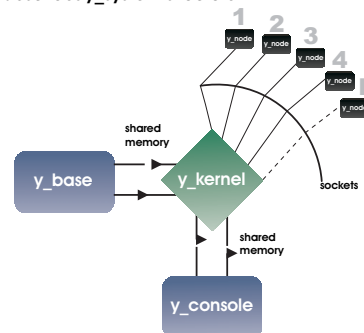


Accuracy dependence of molecular docking is a function of electrostatic coefficient (for molecular graph model).

The best correlation index for molecular graph model is 3.643e-2 (point 2.4), for orbital graph model is 3.169e-2 (point 1.0) and for AM1 semi empirical approach is 3.624e-2 (point 0.6)

## SOFTWARE BACKGROUND

Implementation of orbital graph model for partial atom charges definition was developed as a part of **y\_system**. **y\_system** is a developing project of full flexible receptor based virtual screening system which is directed to distributed computing. We had started this project in autumn of 2004. The core of **y\_system** is an interaction between built-in database engine with several different calculation plug-ins. The picture below describes **y\_system** structure:



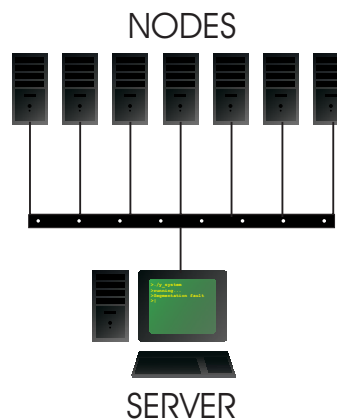
**y\_system** is written in C for Linux OS. It's main features are:

- Integrated databases for the access to molecular complexes coordinates, compound structures, QSAR tables etc
- Distributed calculations support
- Compounds preprocessing
- Entropy-based scoring function for receptor-based virtual screening
- 3D/4D QSAR for virtual complexes sets
- Full-flexible molecular docking
- Protein-protein screening and design support

● Finished ● Under development ● Will be developed in future

## TECHNICAL FACILITIES

Our laboratory has a cluster of 8 computers (P4 3.2GHz, 1 Gb of memory, Linux OS) for the computations. This cluster mainly works under control of **y\_system** program suite. Server node keeps database server **y\_base** and controls all other nodes in cluster via **y\_kernel** and **y\_console**. In addition, every node (including server) has a calculation plug-ins connected via sockets to **y\_kernel** (task manager of **y\_system**).



These hardware are enough for processing about 1 000 000 compounds per month via receptor based virtual screening.

We are interesting in collaboration in the fields of virtual screening of proteins, small compounds as well as modeling algorithms development.

## REFERENCES

1. Oliferenko A.A., Krylenko P.V., Palyulin V.A. and Zefirov N.S. A new scheme for electronegativity equalization as a source of electronic descriptors: application to chemical reactivity. *SAR and QSAR in Environmental Research*, Vol. 13 (2), (2002) 297305