C-7

Development of Multistep Ligand-Based Virtual Screening Cascade Methodology in a Search for Novel HIV-1 Integrase Inhibitors: 2. Privileged Fragments.

Dawid Warszycki [1], *Agata Kurczyk* [2], *Rafał Kafel* [1], *Robert Musioł* [2], *Andrzej J. Bojarski* [1], *Jarosław Polański* [2]

[1] *Department of Medicinal Chemistry, Institute of Pharmacology, Polish Academy of Sciences, Smetna 12, Kraków, Poland*
[2] *Department of Organic Chemistry, Institute of Chemistry, University of Silesia, Szkolna 9, Katowice, Poland*
e-mail: warszyc@if-pan.krakow.pl

HIV integrase which is essential in the virus replication cycle and has no homologue among human enzymes [1], became an important target for drug development more than twenty years ago. Nevertheless, progress has been hampered by the lack of assays suitable for high throughput screening. Thus, a real breakthrough was only observed in 2007 with the introduction of the first integrase inhibitor, raltegravir, into treatment.
Crystal structure for HIV-1 integrase is already known and thus, both techniques commonly used in VS campaigns (structure and ligand-based) could be developed. Here we introduced a multistep ligand-based screening cascade because it is suggested that ligand-based methods outperform structure-based in true positives identification [2]. Our strategy consists of two sequential modules: machine learning-based (ML-based) and privileged fragments-based (PF-based).
The PF module is a weight-based scoring function which rates presence of particular molecular fragments, previously recognized as privileged, in screened compounds. Mentioned fragments are defined as structural subunits specially effective in distinguishing active compounds from inactives. PFs were extracted by using MI-DSE formalism [3] on thirteen unique training sets. Finally, the prepared module was applied as standalone or as a second-step in a multistep VS experiment. The test set was composed from 450 actives (not used for the module development) and 16200 DUD [4] decoys generated by an in-house script. The developed module achieved AUC = 0.760 and more than 40-fold enrichment in a single-step experiment. Overall, two-step VS campaign, where PF-based module was a second-step, reached more than 200-fold enrichment of actives/inactives ratio.

[1] Delelis O., Carayon K., Saib A., Deprez E., Mouscadet J.-F.: *Retrovirology* **5** (2008), 114.
[2] Meslamani J., Li J., Sutter J., Stevens A., Bertrand H.-O., Rognan D.: *J. Chem. Inf. Model.* **52** (2012), 943-955.
[3] Wassermann A.M., Nisius B., Vogt M., Bajorath J.: *J. Chem. Inf. Model*. **50** (2010), 1935-1940.
[4] Huang N., Shoichet B.K., Irwin J.J.: *J. Med. Chem.* **49** (2006), 6789-6801.

This presentation is a continuation of poster entitled ‘Developed of multistep ligand-based virtual screening cascade methodology in a search for novel HIV-1 integrase inhibitors: 1. Machine learning’.