[P37] Structure motifs in HIV1 RT ligand data published in the literature

Birgit Viira, Alfonso Tlatoani Garcia-Sosa, Uko Maran

Institute of Chemistry, University of Tartu, Ravila 14A, 50411, Tartu, Estonia.

The growing interest in HIV1 RT inhibitors and drug-resistant mutations over the past years has led to an increasing amount of data regarding chemical and the corresponding biological activity space. This inhibitors data serves ground for chemoinformatic analysis to understand structural pattern of known active chemicals and most prevailing chemotypes. In the present work, the different kind of chemotypes, which have experimental data against HIV, were investigated.

HIV1 RT inhibitors were extracted from the ChEMBL database (version 18). Database query resulted in 19640 bioactivity ligands. After extensive curating, final dataset consisted of 750 compounds with Ki and Kd values, which were measured against wild type and 13 different HIV1 RT mutants: K103N, L1001I, Y181C, V106A, Y188L, Y181I, M184V, G190A, V179D, K65R, P236L, P119S, T165A. Curated data was analyzed with methods implemented in Scaffold Hunter [http://scaffoldhunter.sourceforge.net], which provides integrated visualization and analysis of biological activity data. The scaffold tree algorithm was used for the hierarchical classification of chemical compound sets based on their common core structures, ie. scaffolds. Analyzing these virtual scaffolds may discover the 'holes' not covered by the compounds in the database and are promising starting points for further investigation.

Altogether six different 'parent' groups were discovered: carbocycles, N-heterocycles, O-heterocycles, N,O-heterocycles, N,S-heterocycles and S-heterocycles. The majority of the inhibitor scaffolds contain at least two rings. The most widespread 'parent' types are N-heterocylces and N,O-heterocycles. To date, 13 approved drugs for HIV1 RT are used: abacavir, delavirdine, didanosine, efavirenz, emtricitabine, etravirine, lamivudine, nevirapine, rilipivirine hydrochloride, stavudine, tenofovir disoproxil, zalcitabine and zidovudine. A scaffold tree was also constructer for approved drugs to analyze their structures and compare them to the previous dataset. 12 molecules out of 13 consist of N-heterocycles and 1 molecule contains a N,O-heterocycle as a 'parent' structure. The approved drugs scaffold tree coincides with the above mentioned data scaffold tree.