INTEGRATED LIGAND AND STRUCTURE BASED APPROACH TO THE SEARCH OF THE HIV-1 REVERSE TRANSCRIPTASE INHIBITORS

O.A. Tarasova, D.A. Filimonov, V.V. Poroikov

Institute of Biomedical Chemistry, Moscow, Russia

The design of new antiretroviral drugs is still of high interest due to the issues of safety and effectiveness of the medicines, which are currently used in clinical practice. HIV RT inhibitors, targeting early stages of the virus-host interaction, are of great interest for the investigators. Now, the antiretroviral drugs that inhibit HIV RT only allow decreasing the HIV replication but does not provide the full elimination of the virus [1]; thus, the studies on new antiretroviral drugs are in high demand.

An acquired HIV RT resistance happens due to the high rate of mutations in the particular region of the pol gene, encoding the HIV RT amino acid sequences [2]. There are a lot of clinical and biochemical data on the relationships between the occurring of the single point mutations and their combinations in the pol and the resistance of the particular variants of the RT to the nucleoside reverse transcriptase inhibitors (NRTI) and non-nucleoside reverse transcriptase inhibitors (NNRTI).

We propose an application of earlier developed PASS algorithm [3] to the (1) prediction of the amino acid changes, potentially involving in the resistance of HIV-1 and (2) integrated ligand and structure-based approach to the search for the compounds with the high potential against the variety of the HIV resistant strains. In our study, we used more than 3200 variants of the HIV-1 RT variants from the publicly accessible HIV Drug Resistance Database [4] tested against the ten anti-HIV drugs including both NRTI and NNRTI (DR data set). The set of HIV-1 RT variants was tested in two different genotype-phenotype test systems: about 1300 HIV-1 RT variants were tested in PhenoSense yielded Phenosense data set and about 1900 variants were tested in Antivirogram (Antivirogram data set). The amino acid position and the particular amino acid itself were used as the descriptors for the prediction. Two classes of the variants were considered: “susceptible” and “resistant”. The average balanced accuracy of prediction obtained in the leave-one-out procedure for the Phenosense data set was about 82%, and for the Antivirogram data set was about 87%.

For further computational experiments, we used DR data set and carefully annotated data set of over 230 variants of the HIV-1 RT inhibitors collected from the Protein Data Bank and NCBI Protein databases. We have developed and tested the novel approach based on the integration of the two types of descriptors. We took into account (1) the conditional probability of the specific pentapeptide to occur in the amino acid sequence of the particular variant of HIV RT and (2) the conditional probability of the ligand descriptors to arise in the particular ligand. In our approach we use two types of descriptors: (1) multilevel neighborhoods of atoms, MNA [3] and (2) quantitative neighborhoods of atoms, QNA [5]. The details of the algorithm and the results of the method validation will be presented.

1. Geronikaki A. et al., *Topics in Medicinal Chemistry*, 1-59. DOI:10.1007/7355_2015_5001.
2. Li W, et al., *Eur J Med Chem.*, 2015, 102, 167-179.
3. Filimonov D. et al., *Chemistry of Heterocyclic Compounds*, 50 (3), 444-457.
4. Shafer R.W., *Journal of Infectious Diseases*, 194 Suppl 1:S51-8.
5. Filimonov D.A. et al., SAR and QSAR in Environmental Research, 20 (7-8), 679-709.