UDC 608.2

AN APPLICATION FOR THE ANALYSIS OF PHARMACOLOGICAL ACTIVITY OF CHEMICAL COMPOUNDS USING CHEMINFORMATICS METHODS

Isakova A.M. (Saint Petersburg State University), **Kovalenko A.A.** (The University of Manchester)
Supervisor – Dr. Shityakov S.V.
(ITMO University)

Introduction. The search for new biologically active molecules is one of the most important tasks of modern chemistry and biotechnology. Traditionally, chemists synthesize a new molecule in a laboratory and screen it, i.e. determine whether this molecule actually has the appropriate properties. Such studies require a lot of time and money. Only a small part of the created active substances successfully passes the screening stage, but capitalized clinical costs were estimated to be from $291 million to $412 million. It takes almost 12-15 years to develop a single medicinal product from the time the new molecule is discovered to the time when the product is available in market for treating patients. So, software for the first stage screening of potential drug molecules is needed [1]. At the moment, all of the existing solutions are either outdated and difficult to run on modern devices, or inconvenient to use and provide little functionality.

Main part. The aim of our project was to develop software that, having received encoded representations of reagents as input, will produce appropriate encoded representations of reaction products and analyze their pharmacological activity. In the first stage, we need to upload reagent molecules. The molecules are represented in SMILES [2] format because it is supported by the most chemical databases and the most convenient for pharmacological descriptors calculations. Next, the reaction process is launched and the program shows some products generation statistics. We use SMARTS patterns to describe the reaction mechanism. The user has the ability to filter the reaction products by 34 pharmacological descriptors [3], including physical properties, pharmacokinetics descriptors, logP, logS, Lipinski rule and so on. The user can save program results in 39 different file formats, including .txt, .csv, .mol, .fasta, .sdf, .pdb. It gives an opportunity to do a more detailed further analysis using other computational methods (i.e. molecular docking, molecular dynamics, etc.) and to visualize chemical compounds in 2D or 3D.

Conclusion. We developed the unique software with a high performance, processing datasets from thousands-to-millions compounds in minutes on a desktop computer. Our application is adapted to all the most popular modern operating systems (Windows, Linux, MacOS) and can be used as a standalone version.

Literature review :

1. Nigam, A. K., Pollice, R., Hurley, M. F. D., Hickman, R. J., Aldeghi, M., Yoshikawa, N., Chithrananda, S., Voelz, V. A., & Aspuru-Guzik, A. Assigning confidence to molecular property prediction // Expert Opinion on Drug Discovery. – 2021. – V. 16. – № 9. – P. 1009–1023.
2. Weininger, D., Weininger, A., & Weininger, J. L. SMILES. 2. Algorithm for Generation of Unique SMILES Notation//Journal of Chemical Information and Computer Sciences. – 1989. – V. 29. – № 2. – P. 97–101.
3. Daina, A., Michielin, O., & Zoete, V. (2017). SwissADME: A free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. *Scientific Reports*, *7*. https://doi.org/10.1038/srep42717