DOI: 10.46991/JISEES.2025.SI1.184

Discovering Drug Leads by Ultrafast Docking Screens in Large Chemical Spaces and AI/ML Pipeline

Ruben Abagyan[1*], Max Totrov[2], Hovakim Grabski[3], Siranuysh Grabska[3,4]

[1] University of California San Diego, La Jolla, CA, USA
[2] MolSoft, LLC, San Diego, CA, USA
[3] L.A. Orbeli Institute of Physiology, National Academy of Sciences, Yerevan, Armenia
[4] Microbiology & Biotechnology, Yerevan State University, Yerevan, Armenia

ABSTRACT

Accelerated discovery of potent and selective inhibitors for challenging emerging targets holds promise for many kinds of incurable diseases including certain stages of types of cancer, autoimmune diseases, neurodegenerative diseases, and rare and neglected diseases. Identifying the first potent and selective inhibitors for transient, allosteric, or protein–protein interaction (PPI) pockets remain a significant challenge in drug discovery. These targets often require an integrated approach that combines structural biology, multi-modal data, sophisticated computational tools, and ultra-large-scale virtual screening coupled with in silico compound optimization. Key challenges include:

1. Characterizing the molecular target, including its conformational and mutational states, and identifying a druggable binding pocket;
2. Discovering novel chemical entities with nanomolar potency against the selected target;
3. Eliminating candidates with undesirable off-target activities;
4. Optimizing ADMETox (absorption, distribution, metabolism, excretion, and toxicity) properties and achieving the required tissue and cellular distribution.

To address these challenges, we have developed a high-performance CPU/GPU-accelerated computational pipeline in collaboration with Molsoft, and deployed it at the UCSD laboratory and in Armenia at Orbeli PI, YSU/YerevaNN and Center for Sci.Comp. This platform enables the screening of billions of compounds and supports expansive generative and combinatorial exploration of chemical space. We present case studies where this pipeline successfully identified drug candidates in collaborations with biotechnology startups, with some candidates already advanced toward clinical development and US trials. These examples highlight the platform's potential to transform the early stages of drug discovery for difficult and previously intractable targets.

Keywords: structure-based drug discovery, docking, screening of billion-sized chemical spaces, preclinical cancer drugs

References:

1. Raush, E.; Abagyan, R.; Totrov, M. Efficient generation of conformer ensembles using internal coordinates and a generative directional graph convolution neural network. *J. Chem. Theory Comput.* **2024,** *20*, 4054–63. DOI:10.1021/acs.jctc.4c00280
2. Abagyan, R.; Totrov, M.; Kuznetsov, D. ICM—A new method for protein modeling and design: Applications to docking and structure prediction from the distorted native conformation. *J. Comput. Chem*. **1994**, *15*, 488–506. DOI:10.1002/jcc.540150503

[*]Corresponding Author:

Ruben Abagyan, UC San Diego, San Diego Supercomputer Center, Moores Cancer Center Structural and Functional Genomics Division.
Email: ruben@ucsd.edu