Math approach may make drug discovery more effective, efficient

November 28 2022

Credit: Pixabay/CC0 Public Domain

Researchers at The University of Texas at Dallas and Novartis Pharmaceuticals Corp. have devised a computer-based platform for drug discovery that could make the process more effective, more efficient and less costly.

Dr. Baris Coskunuzer, professor of mathematical sciences at UT Dallas, and his colleagues developed an approach based on topological data analysis to screen thousands of possible drug candidates virtually and narrow the compound candidates considerably to those that are most fit for laboratory and clinical testing.

The researchers will present their findings at the 36th Conference on Neural Information Processing Systems, which will be held Nov. 28 through Dec. 9 in New Orleans.

Typically, the early phases of drug discovery involve researchers identifying a biological target, such as a protein associated with a disease of interest. The next step is to screen libraries of thousands of potential chemical compounds that might be effective or could be modified to affect the target to alleviate the disease's cause or symptoms. The most promising candidates move on to the lengthy and expensive process of laboratory and clinical testing and regulatory approval.

"The drug-discovery process can take 10 to 15 years and cost a billion dollars," Coskunuzer said. "Drug companies want a more cost-effective way to do this. They want to find the most promising compounds at the beginning of the process so they're not wasting time testing dead ends. We have provided a completely new method of virtual screening that is computationally efficient and ranks compounds based on how likely they

are to work."

While virtual screening of libraries of chemical compounds is not new, Coskunuzer said his group's approach significantly outperforms other state-of-the-art methods on large data sets.

The UTD and Novartis team framed the virtual screening process as a new type of topology-based graph ranking problem, from a branch of mathematics called topological data analysis. Their method characterizes each molecular compound based on the shape of its underlying physical substructure—its topology—as well as a series of physical and chemical properties of the components of the molecule. From this information, the researchers develop a unique "topological fingerprint" for each compound that is used to rank it according to how well it fits the desired properties.

"The advantage of our algorithm is that it could screen about 100,000 compounds in a couple of days, which is much faster than other methods," Coskunuzer said.

The next step will be to generalize the method to molecular property prediction, which includes scoring a compound on how soluble it is in water. Solubility can be critical to a drug's efficacy in the human body.

"If you find a good compound, but it does not have the desired molecular properties—if it's not soluble—then it's likely that it is not going to work. You want to be able to test these properties first before a drug candidate gets too far into development," Coskunuzer said.

More information: Andac Demir et al, ToDD: Topological Compound Fingerprinting in Computer-Aided Drug Discovery, (2022). DOI: 10.1101/2022.11.08.515685

36th Conference on Neural Information Processing Systems

Provided by University of Texas at Dallas