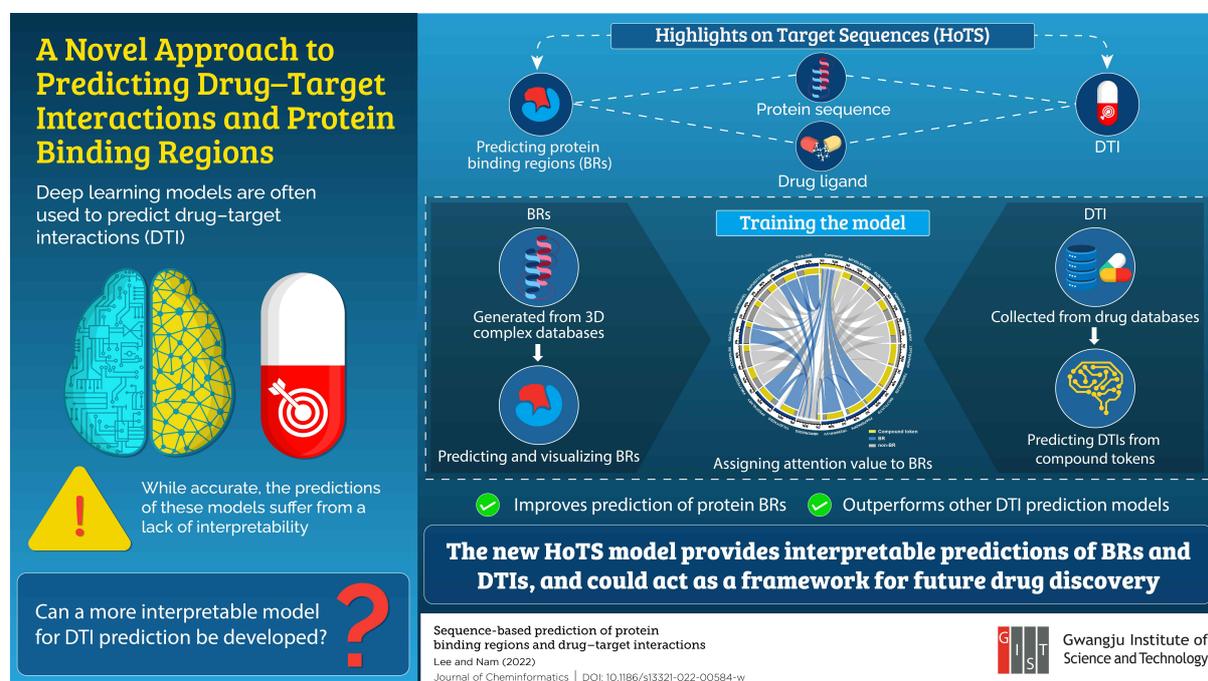


PRESS RELEASE

Researchers at the Gwangju Institute of Science and Technology Identify New Medicines Using Interpretable Deep Learning Predictions

New drugs are predicted for target proteins using a new, interpretable deep learning-based model with improved prediction and transparency

In the drug discovery process, drugs are tested for their ability to bind or interact with target molecules in the body. While deep learning models have made this process quite efficient, their predictions often lack interpretability. Addressing this, researchers at the Gwangju Institute of Science and Technology develop a new deep learning model that not only makes better predictions of drug-target interactions but does so interpretably, providing a good starting point for further drug discovery.



An important aspect of drug discovery involves determining how well a drug binds to its target (protein) molecule. Typically, this step involves aligning a 3D structure of a drug and its target protein at various configurations in a process called “docking.” Preferred binding sites are then discovered by running docking simulations repeatedly with multiple drug candidates for a particular target molecule. Typically, owing to the vast number of likely candidates for potential drug-target interactions, deep learning models are used to carry out such simulations. However, a problem with using such models is the difficulty in interpreting their predictions. While deep learning certainly makes for a speedy drug discovery, it is little more than a black box. Additionally, docking simulations cannot be used to develop drugs for novel targets that have no known 3D complex with an interacting drug.

[In a recent study published in the *Journal of Cheminformatics*](#), Associate professor Hojung Nam and her Ph.D. student Ingoo Lee from the Gwangju Institute of Science and Technology

in Korea developed an new deep learning model called “Highlights on Target Sequences” (HoTS) for predicting the binding between a drug and a target molecule. The new model not only makes better predictions but, more interestingly, did so in an interpretable manner. Moreover, the model could predict drug-target interactions (DTIs) without the need for simulations or 3D structures.

How did the team achieve this remarkable feat? Professor Nam explains how their drug-target interaction prediction model works: *“First, we explicitly teach the model which parts of a protein sequence will interact with the drug using prior knowledge. The trained model is then utilized to recognize and predict interactions between drugs and target proteins, giving better prediction performances. Using this, we built a model that can predict the target proteins' binding regions and their interactions with drugs without a 3D-complex.”*

Rather than dealing with the complete length of the protein sequence, the model could make predictions based only on the parts of the protein that are relevant to the DTI interaction. *“We taught the model where to ‘focus’ to ensure that it can comprehend important sub-regions of proteins in predicting its interaction with candidate drugs,”* elaborates Professor Nam. This, in turn, allowed the model to predict DTIs more accurately than existing models.

Considering that the model does not require information about the 3D structure and provides predictions that are clearly understandable, the findings of this study provide a good starting point for further docking simulations to predict new drug candidates for targeted therapy, accelerating advancements in this field. *“The model used in our study would make the drug discovery process more transparent as well as low-risk and low-cost. This will allow researchers to discover more drugs for the same amount of budget and time,”* concludes Professor Nam.

Let’s hope her visions are not too far from being realized!

Reference

Authors:	Ingoo Lee, Hojung Nam*
Title of original paper:	Sequence-based prediction of protein binding regions and drug–target interactions
Journal:	<i>Journal of Cheminformatics</i>
DOI:	10.1186/s13321-022-00584-w
Affiliations:	Gwangju Institute of Science and Technology, Republic of Korea

*Corresponding author’s email: hjnam@gist.ac.kr

About the Gwangju Institute of Science and Technology (GIST)

The Gwangju Institute of Science and Technology (GIST) was founded in 1993 by the Korean government as a research-oriented graduate school to help ensure Korea's continued

economic growth and prosperity by developing advanced science and technology with an emphasis on collaboration with the international community. Since that time, GIST has pioneered a highly regarded undergraduate science curriculum in 2010 that has become a model for other science universities in Korea. To learn more about GIST and its exciting opportunities for researchers and students alike, please visit: <http://www.gist.ac.kr/>.

About the author

Hojung Nam is an Associate Professor in the School of Electrical Engineering and Computer Science at the Gwangju Institute of Science and Technology (GIST). Her research group is developing various computational tools for drug discovery and development. For example, her group created a computational model to predict side effects caused by drug-drug interactions, predicting drug responses for patients, and developed generative models for novel molecules. Before coming to GIST, she completed her postdoctoral training in the Department of Bioengineering at University of California in San Diego (UCSD). In 2009, Hojung Nam received a Ph.D. in the Department of Bio and Brain Engineering at the Korean Advanced Institute of Science and Technology (KAIST).

Ingoo Lee is a Ph.D. candidate under her advisor Professor Hojung Nam at GIST. During his Ph.D. training, he has shown great progress in building drug-target interactions prediction models.