[BIOINFO 2024]

Reinforcement Learning-based Translational Fusion Partner Engineering

Seongbo Heo[1, 2], Bong Hyun Sung[1, 2], Jung-Hoon Sohn[1, 2], Dae-Hee LEE[1, 2], Seung-Goo LEE[1, 3],

Haseong Kim[1, 2, *]

[1] *Department of Biosystems and Bioengineering, Korea National University of Science and Technology*
[2] *Synthetic Biology and Bioengineering Research Center, Korea Research Institute of Bioscience and Biotechnology*
[3] *Synthetic Biology and Bioengineering Institute, Korea Research Institute of Bioscience and Biotechnology*
**Corresponding author: haseong@kribb.re.kr*

To enhance the secretion efficiency of target proteins in yeast which is widely used strain for protein production, Translational Fusion Partners (TFPs) are employed. However, it is difficult to design optimal TFP sequences due to biochemical complexity with its target protein. In this study, we developed an artificial intelligence model based on reinforcement learning to automatically generate optimal TFP sequences. We firstly developed a protein sequence-based model that predicts relative titers of secreted target proteins using a protein large language model. Then the model was used as an environment for the feedback of a TFP sequence generation model during reinforcement learning. As a result, the titer prediction model achieved acceptable performance with low errors. We also confirmed that TFP sequences generated through the sequence generation model exhibited high sequence similarity to sequences that actually demonstrate high relative secretion levels. This study is expected to make a significant contribution to the field of protein engineering by presenting an effective method for designing TFPs to optimize protein secretion in yeast, and by applying it to various protein production systems, it is anticipated to broaden its industrial applicability.