세미나 초록

성명	김동혁
소속	울산과학기술원
발표 주제	Transformer-based analysis pipeline for classification and engineering of enzymes
발표 내용	Understanding and manipulating the enzymatic preference of NAD(P) cofactors is increasingly crucial in metabolic engineering. However, despite the wealth of knowledge, designing mutants to switch cofactor specificity remains as a complex task. Traditional structure-guided methods are knowledge-intensive, while bioinformatic approaches are often limited by structural and computational factors. Here, we introduce DISCODE (Deep learning-based iterative pipeline to analyze Specificity of Cofactors and to Design Enzyme), a novel transformer-based deep learning model to predict NAD(P) cofactor preferences. For model training, a total of 7,132 NAD(P)-dependent enzyme sequences were collected. Leveraging the transformer architecture, DISCODE classifies the cofactor preferences of NAD(P)-dependent oxidoreductases protein sequences without structural or taxonomic limitation. The model showed 97.4% and 97.3% of accuracy and F1 score, respectively. A notable feature of DISCODE is the interpretability of its attention layers of the transformer-based deep learning model, which facilitates the identification of key residues for cofactor specificity. The residue selection guided by attention analysis was highly consistent with experimentally verified designs in cofactor switching, expanding the model's practical utility. Integrated into an enzyme design pipeline, DISCODE coupled with attention analysis, enables a fully automated approach to redesign cofactor specificity.