



Contribution ID: 9

Type: **Talk**

AI-powered protein engineering: the next frontier in biocatalyst development

Tuesday 20 February 2024 15:15 (30 minutes)

Enzymes are a transformative tool for the path towards a digital bioeconomy due to their ability to catalyze diverse biochemical processes in sustainable and adaptable ways. Protein engineering plays a pivotal role in advancing the development of such biocatalysts with applications spanning biotechnology, biomedicine, and life sciences. The advent of protein-directed evolution, recognized with the 2018 Nobel Prize, has enabled the customization of enzyme functions for a broad range of new-to-nature applications. Despite significant progress, challenges persist due to a limited understanding of protein function and the complex multi-factorial optimization problem inherent in enzymes and proteins. Recent advancements, however, integrate machine learning (ML) techniques to address these challenges. Exazyme focuses on leveraging ML algorithms to enhance protein functionalities, including enzyme properties.

Our methodology employs a two-step ML model application. Initially, our models proficiently predict protein sequence-to-function mappings based on functionally assayed sequence variants, requiring minimal reliance on detailed mechanistic or structural data. This approach has proven particularly useful in low data regimes, where extensive screening is not possible. Subsequently, we utilize these predictions in a Bayesian optimization framework to guide the selection of candidates for experimental validation, enabling simultaneous optimization of multiple parameters, such as stability, catalytic speed, and substrate specificity. A noteworthy accomplishment of our research lies in the superior performance of our prediction algorithms, consistently outperforming current state-of-the-art methods, across various datasets and benchmarks. Practical validation of our algorithms was demonstrated through successful protein engineering campaigns, enhancing the functionality of complex enzymes from diverse families including carboxylases, hydrogenases, and phosphohydrolases.

Our results underscore the capacity of ML methods to expedite the processes of directed evolution and rational design of proteins. By efficiently predicting and selectively identifying sequences with improved properties, these methods leverage existing sequence variant data to advance the field of protein engineering, and thereby a digital bioeconomy.

Authors

Affiliation

Consent

Presenter: IVANOVSKA, Jelena (Exazyme GmbH)