

AI-empowered Precise Structural Regulation and Multidimensional Directed Evolution of Industrial Enzymes

Yi Zhao¹, Mingjia Yu^{2, *}

¹Frontier Interdisciplinary Domain, Beijing Institute of Technology (Zhuhai), Zhuhai, China

²School of Chemistry and Chemical Engineering, Beijing Institute of Technology, Beijing, China

Email address:

3120256737@bitzh.edu.cn (Yi Zhao), mjyu@bit.edu.cn (Mingjia Yu)

*Corresponding author

Abstract

The directed evolution of industrial enzymes holds significant strategic value in green biomanufacturing. However, traditional engineering methods face critical bottlenecks. These include the vast sequence space and the challenge of synergistically optimizing multidimensional properties such as activity and stability. To overcome these limitations, we developed a highly efficient closed-loop system integrating rational protein design with artificial intelligence (AI)-assisted computation. To elucidate poorly understood structure-function relationships, we combined molecular docking, molecular dynamics simulations, and density functional theory (DFT) calculations. By incorporating deep learning tools such as RFDiffusion, we accurately characterized the catalytic mechanisms and active site conformations of complex enzymes including 3-OST and TbSMT. Furthermore, to address the low efficiency of monomeric stability engineering, we established a multi-energy filtered stability screening strategy (MEFS Pipeline) alongside a high-precision thermal stability prediction model. This pipeline computationally prioritized high-potential mutants and seamlessly bridged *in silico* predictions and wet-lab validations. It directly guided downstream experiments including protein expression, purification, and catalytic kinetics assays. For instance, the W59R mutant of 3-OST was designed by remodeling its spatial electrostatic distribution. This mutation significantly increased the overall electrostatic potential of the protein from 9,109.99 kT/e to 24,320.46 kT/e. Consequently, it enhanced the long-range electrostatic guidance and substrate capture of negatively charged polysaccharides. Additionally, our thermal stability prediction model (SMRC-Net) achieved an R^2 of 0.89 and an RMSE of 0.79. It also demonstrated 90% accuracy in predicting the direction of stability changes for independently designed mutants. To further enable the synergistic evolution of multiple enzyme attributes, we constructed a multimodal machine learning framework. This framework integrated pre-trained protein language models with topological descriptors such as persistent Laplacians. In turn, it successfully facilitated the co-optimization of protein solubility and thermal stability. Tightly coupling computational prescreening with rigorous wet-lab experimentation drastically reduces the costs associated with blind trial and error. Ultimately, this study demonstrates that AI-driven computational design substantially empowers rational macromolecular engineering. It provides a reliable framework to accelerate the transition of enzymes from natural states to industrially customized functions.

Keywords

Directed Evolution of Enzymes, Rational Protein Design, AI-assisted Computation, Multidimensional Property Optimization, Conformational Stability