XenoFold AI: A Novel Application for Chemistry-Led, AI-Powered Candidate Enzyme Design for the Xenobiotic Degradation of Pollutants

Torres, Kinsley (School: Perry High School)

XenoFold is a chemistry-conditioned generative framework designed to computationally propose enzymes capable of degrading xenobiotic pollutants. The system integrates reaction-aware molecular encoding with protein sequence generation to produce candidate enzymes tailored to specific chemical transformations. Reaction representations derived from SMILES strings are embedded using a molecular transformer encoder and used to condition a sequence-to-sequence protein language model that generates novel amino acid sequences predicted to catalyze the specified reaction. Generated sequences are evaluated through a multi-stage in silico validation pipeline that includes structural prediction using ESMFold, catalytic plausibility filtering based on predicted residue geometry, and molecular docking with AutoDock Vina to assess substrate binding. Additional metrics such as predicted folding confidence (pLDDT), ligand efficiency, and estimated kinetic parameters are used to prioritize candidates. The framework is designed to generate extracellular, cofactor-independent enzymes targeting environmentally persistent compounds including plasticizers, flame retardants, and halogenated industrial pollutants. By conditioning protein generation directly on chemical reactions rather than known enzyme templates, XenoFold explores a broader functional sequence space for potential biocatalysts. This approach demonstrates the potential of reaction-conditioned generative models to accelerate early-stage enzyme discovery for environmental remediation.

Awards Won:

Non-Trivial Ventures: Non-Trivial Fellowship Scholarship