Zero-Shot Model-Guided Engineering of a Fluorinase via Single Point Mutations

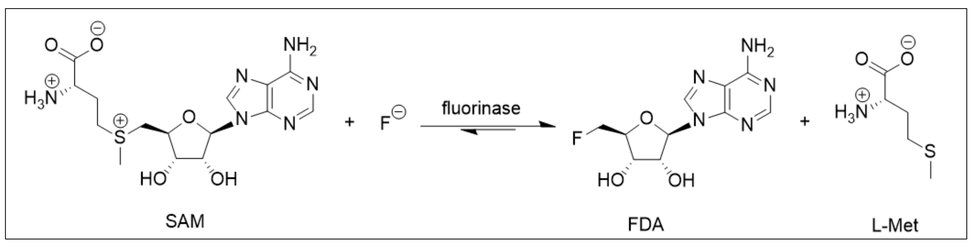
David Harding-Larsen[1], Brianna M. Lax[1], Christopher K. Weingarten[1], Aboubakar S. Sako[1], Stanislav Mazurenko[2,3], Ditte H. Welner[1]

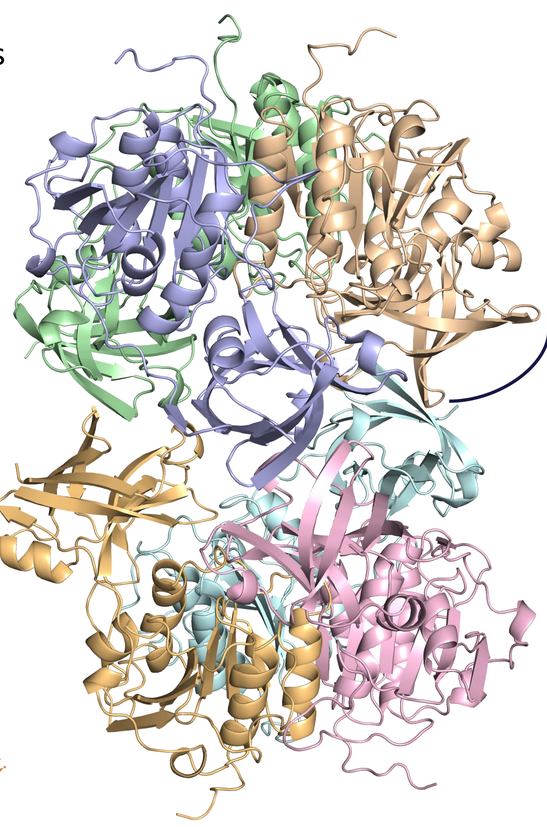
[1] Novo Nordisk Foundation Center for Biosustainability at the Technical University of Denmark, [2] Loschmidt Laboratories of Masaryk University,
[3] International Clinical Research Center, St. Anne's University Hospital Brno, Czech Republic

1) The Fluorinase

Background and motivation

- Many compounds used in agriculture[1] and pharmaceuticals[2] contain carbon-fluoride (C-F) bonds
- The fluorinase is the only known enzyme to catalyze C-F bond formation[3], converting S-adenosyl methionine (SAM) to 5-Fluorodeoxyadenosine (FDA)
- The fluorinase has a very slow turnover rate (k_{cat} = 0.2 min^{-1})[4], limiting industrial applications
- Previous engineering attempts have been mostly unsuccessful at increasing the turnover rate[4,5,6]
- The low number of fluorinase homologues and available data limits traditional Machine Learning (ML) approaches

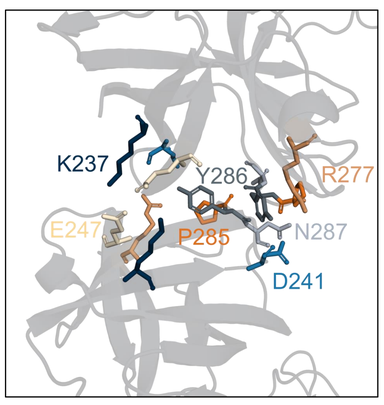


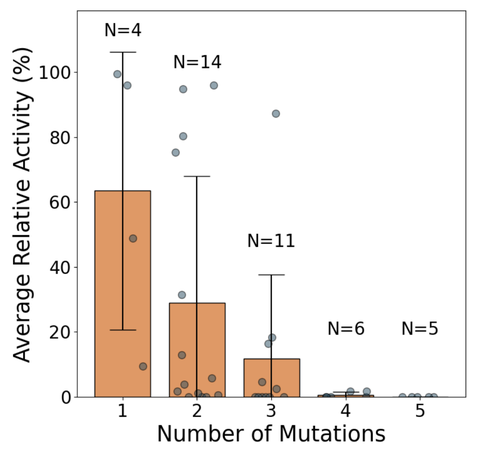


2) Initial Dataset

Mutational scan of hexamer interface residues

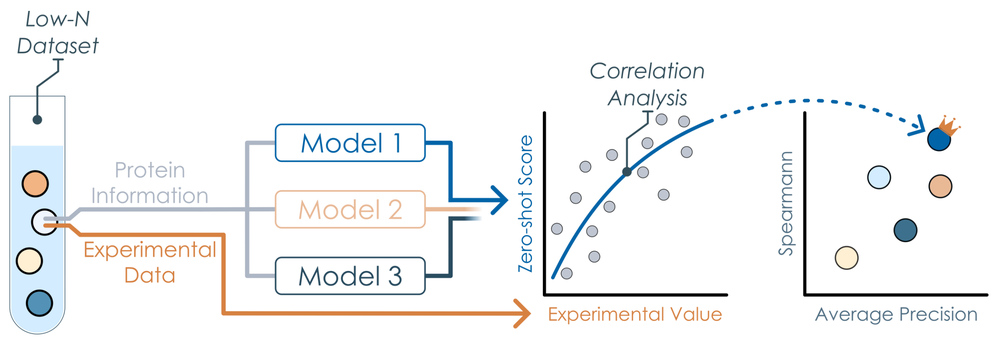
- The fluorinase is a homo-hexamer, with two trimeric assemblies forming a dimer
- The dimerization has been shown to be beneficial, but not necessary, for activity
- A mutational scan was performed on the hexameric interface, including 7 positions and up to 5 mutations
- Unfortunately, no mutants faster than wildtype (WT) could be identified

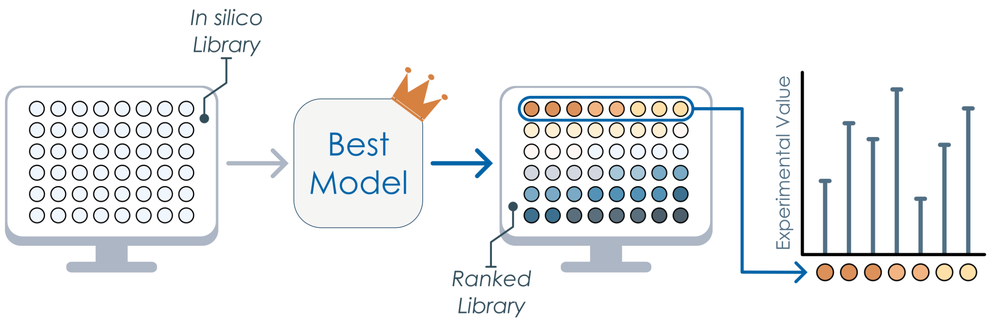




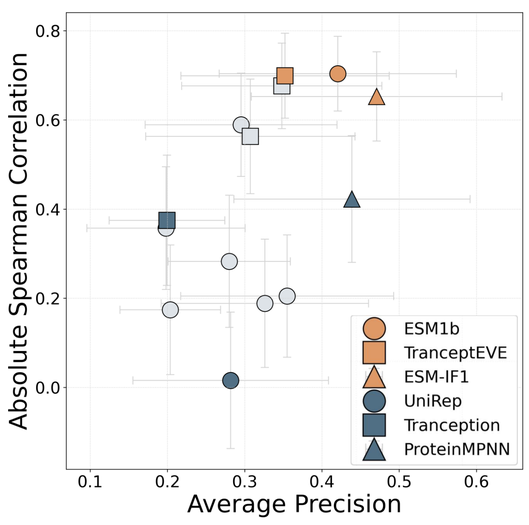
3) PRIZM

Protein Ranking using Informed Zero-shot Modelling





- The mutational scan at the hexameric interface was used as a low-N dataset for PRIZM selection
- The best zero-shot model for each model type (Sequence-, Multiple Sequence Alignment (MSA-), & Structure-based) was identified
- Top single-point mutations predicted to be faster than WT were selected from each model, as well as combinations of the models



Best Models

Worst Models

Sequence

MSA

Structure

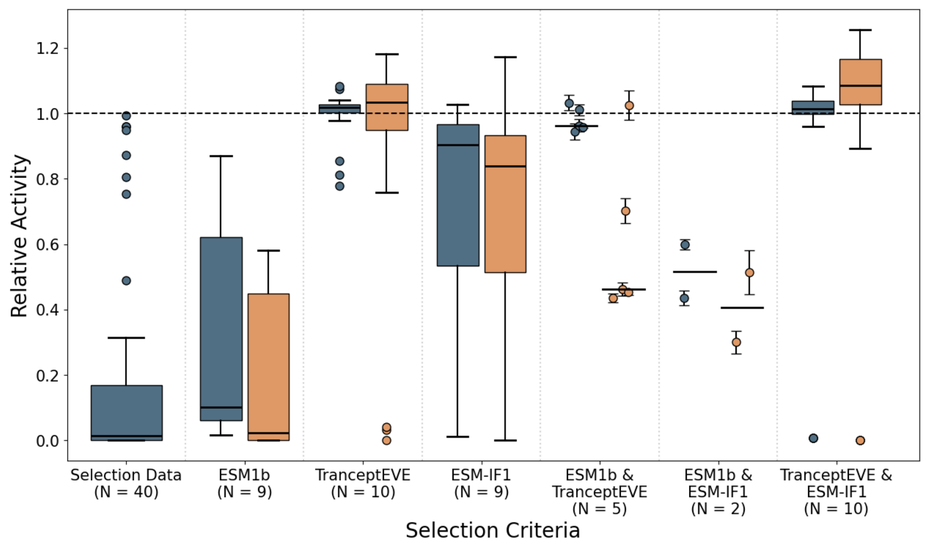
- Best models: ESM1b, TranceptEVE, ESM-IF1
- Top 10 mutants of each model expressed and purified for testing
- Top 10 (or all mutations predicted to be faster if less than 10) of each model combination tested as well
- No mutations agreed upon by all three models

4) Model Results

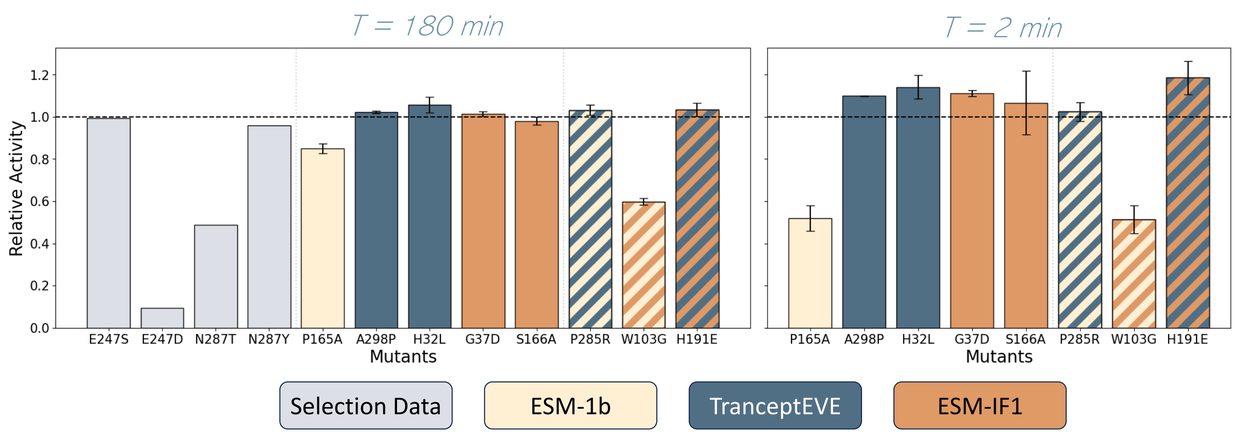
Testing relative activity of predicted mutations

T = 180 min

T = 2 min



- Mutations predicted by the selected models outperform mutations in selection dataset on average
- The contrast between model performance further increases when taking an earlier timepoint than the 3h from the selection dataset
- TranceptEVE is the most promising model, especially in combination with ESM-IF1



Conclusion & next steps

- PRIZM can act as a great starting point for engineering even challenging enzymes like the fluorinase
- Combining of mutations through “greedy” and rational engineering approaches
- Using these results to further understand structure-function relationship of the fluorinase for future engineering efforts

References:
1) Martinelli and Nikel, Microb. Biotechnol. 2019
2) Chandra, Singh, Mahato, and Patel, Chem. Pap. 2023
3) Tsiavaliaris, Fujita-Becker, and Manstein, Nature 2004
4) Sun *et al.*, Wiley-VCH 2016
5) Kittilä *et al.*, Microb. Biotech. 2022
6) Feng, Cao, Liu, and Xian, Front. in Bioeng. and Biotech. 2022

eensb Enzyme Engineering and Structural Biology
novo nordisk fonden Benefiting people and society