Computational implementation of PBS

PBS takes the enzyme-substrate complex structure (i.e., that considered as the ‘‘wild-type’’ structure) and the target mutant library of screening as input and predicts beneficial mutations as output. The PBS workflow consists of four steps, namely mutation, thermostability screening, TS-analog binding screening, and reactivity ranking (Figure 3). These steps operate in an automated manner using our high-throughput enzyme modeling platform EnzyHTP. Below, we detail each step in the workflow.

Step 1: Mutation

The workflow begins with constructing a mutant library from the prepared wild-type enzyme–substrate complex. Using this structure as the reference, candidate variants are generated through site-saturation mutagenesis, random mutagenesis, rational design, or other empirical strategies. EnzyHTP provides the assign_mutant API to automate this process. In this study, site-saturation mutagenesis was employed to define the starting pool of mutants.

Step 2: Thermostability screening

Next, thermally unstable mutants are removed. The stability of each variant is evaluated by the change in folding free energy (ΔΔG_fold) relative to wild type, computed via Rosetta’s cartesian_ddg protocol and automated through the get_rosetta_ddg API in EnzyHTP. Mutants are ranked by ΔΔG_fold, with positive values indicating destabilization. To reduce downstream computational cost, approximately 60% of predicted unstable variants were discarded in this work.

Step 3: TS-analog binding screening

The reduced library is then evaluated for transition-state (TS) analog binding. Mutant–substrate complexes are constructed by replacing target side chains (via EnzyHTP, integrating PyMol and Amber). These serve as inputs for molecular dynamics (MD) simulations, with constraints applied to reaction coordinates (bond distances, angles, etc.) to approximate the pre-reaction complex as a TS analog. After equilibration, 100-ns production trajectories are sampled as conformational ensembles.

For each mutant: Binding enthalpy is estimated with the MM-PBSA method. Active-site flexibility is measured as the RMSD of residues within 5 Å of the substrate. Lower RMSD indicates greater rigidity, which correlates with reduced activation entropy and improved catalysis. Mutants are ranked separately by MMPBSA energy (strong → weak binding) and active-site RMSD (rigid → flexible). Only variants consistently ranking high in both lists are retained. In this study, the top 40 candidates were selected as a practical balance between computational filtering and experimental feasibility.

Step 4: Reactivity ranking

The final step evaluates chemical reactivity using the electrostatic stabilization energy of the TS (ΔΔG_ele). This descriptor reflects how the enzyme's interior electric field stabilizes a reacting dipole, which correlates with changes in activation energy or enthalpy. For each MD snapshot: A QM cluster model of the reacting species and surrounding residues is built.Single-point energy calculations are performed with Gaussian16 via the EnzyHTP interface. The bond dipole is derived using Multiwfn99, and ΔΔG_ele is calculated as the negative dot product between the dipole and the local electric field.

Mutants are ranked from negative to positive ΔΔG_ele, with more negative values indicating stronger TS stabilization relative to wild type. The top ten variants were recommended for experimental validation in this work. Importantly, the final number of selected mutants is case-dependent, determined by balancing experimental resources with the potential benefit of testing additional candidates.