A Computational Framework for Large-Scale Mechanistic Annotation of Enzymatic Reactions

Background/Objective

- Many of the thousands of biological transformations in living systems lack detailed understanding of enzyme mechanisms required for de novo enzyme and pathway design, and existing computational tools require elusive 3D structural data and knowledge of amino acid residues. A predictor that relies solely on reaction stoichiometry could overcome these barriers.

Approach

- MechFind predicts multi-step mechanisms using only reaction stoichiometry as input.
- Reaction steps are abstracted as changes in chemical moieties, using curated rules from the Mechanism and Catalytic Site Atlas (M-CSA) database.
- Results are analyzed and re-ranked to identify the most chemically plausible pathways with the fewest steps.

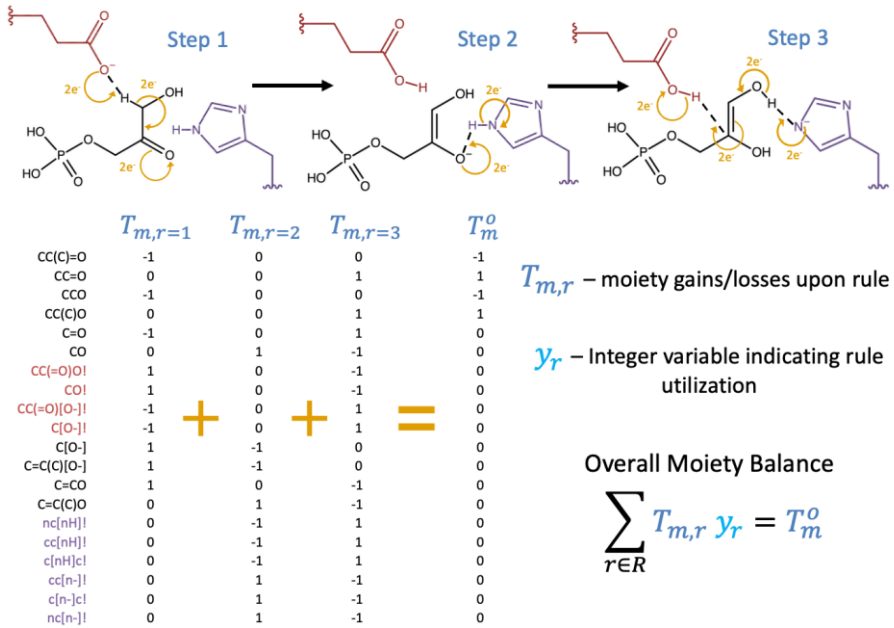
Results

- Mechfind recovered the validated mechanism within the top ten predictions for 85% of the M-CSA dataset and generated plausible mechanistic solutions for nearly 60% of reactions in Rhea and MetaNetX reaction databases.
- Mechfind achieved a 10-fold increase over prior tools, generating over 18,000 hypotheses, and successfully predicted the validated mechanisms for an independent test set of six enzymes.

Significance/Impacts

- By bypassing structural bottlenecks, Mechfind enables the high-throughput annotation of reaction databases and acts as an exploratory engine that maps the landscape of catalytic strategies available for a single reaction.
- Mechfind turns enzyme design into an actionable engineering task by providing information on the potential active site needed by AI protein design models to build more efficient and selective enzymes.

MechFind encodes the mechanism for triosephosphate isomerase from the M-CSA database into three rules that must equate to the overall reaction, enforced by the overall moiety balance constraint.



Performance of MechFind on Rhea and MetaNetX databases.

Database	Rhea	MetaNetX
Reactions tested	14,931	23,112
Predicted mechanisms	8452 (56.6%)	13,658 (59.1%)
Reached 20-min time limit	2539 (17.0%)	4491 (19.4%)
Incompatible reactions	3940 (26.4%)	4963 (21.5%)

Hartley, A.D. et al. *Nat Commun* 17, 3903 (2026). doi: 10.1038/s41467-026-71957-0