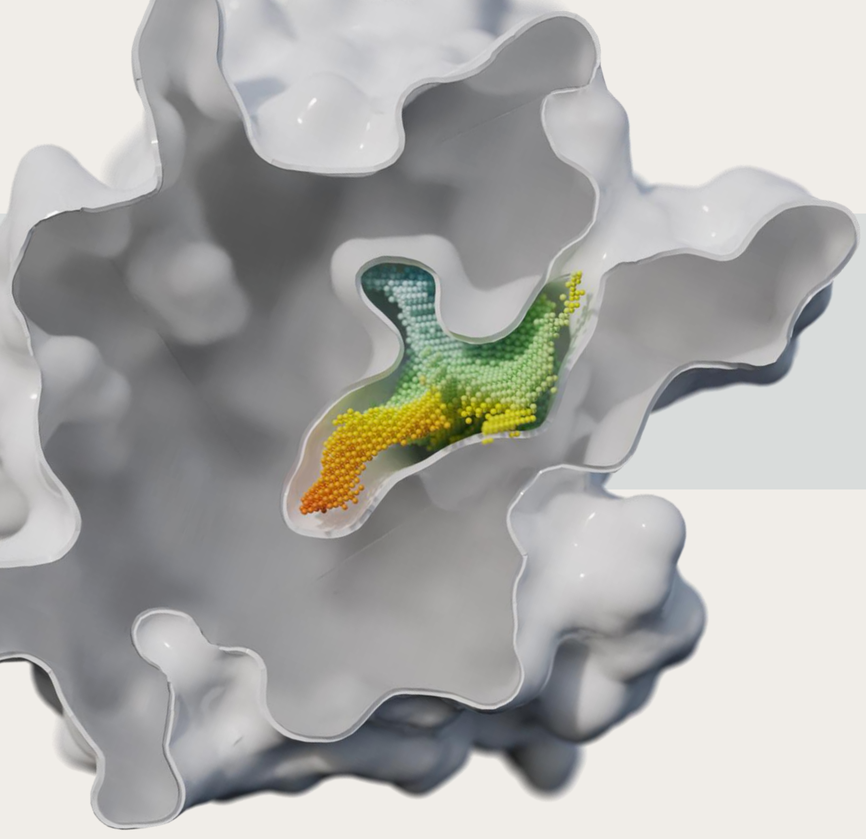


Exploring IREDs with Catalophore-AI

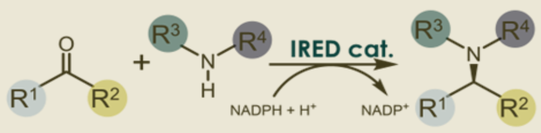
Shifting Frontier for Broad-Scope Reductive Aminations

Bettina M. Nestl[1] , Marco Cespugli[1] , Andreas Krassnigg[1] , Tobias Schopper[1] , Alice Guarneri[1] , Chiara Gasbarri[1] , Georg Steinkellner[1], Christian C. Gruber[1] , Irene Marzuoli[2] , Hans Iding[2] , Serena Bisagni[2] , Kurt Püntener[2] , Francis Gosselin[3] , Sarah A. Berger[4] , Christopher Grimm[4] , Joerg H. Schrittwieser[4] & Wolfgang Kroutil[4]

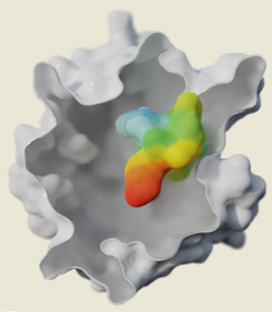
[1]Innophore GmbH, 8010 Graz, Austria
[2]F. Hoffmann-La Roche Ltd., Process Chemistry and Catalysis, 4070 Basel, Switzerland
[3]Genentech Inc., Department of Small Molecule Process Chemistry, 94080 San Francisco, USA
[4]Institute of Chemistry, University of Graz, 8010 Graz, Austria

DESIGNING A BROADLY APPLICABLE WORKFLOW

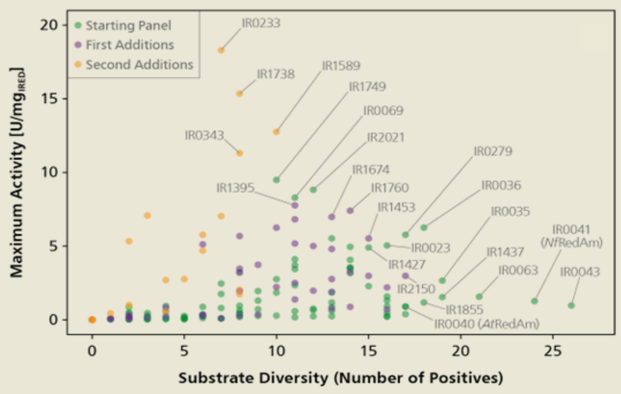
Here, we present an iterative, multidimensional strategy that merges both computational and experimental data to explore the imine reductase[1] (IRED/RedAm) sequence and chemical (substrate scope) space while also predicting the performance of new IREDs as well as substrates.

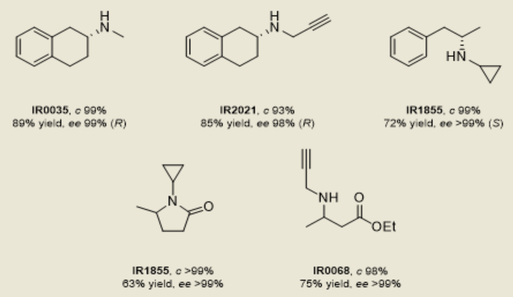


IRED screening against substrates (1:1 eq. 50 mM concentration at pH 8) to test reductive aminase activities.



IREDs features analysis using catalophore 3D-point clouds.



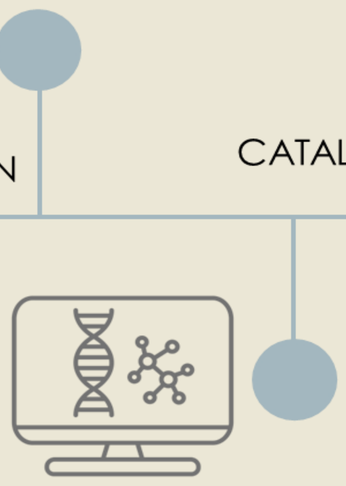


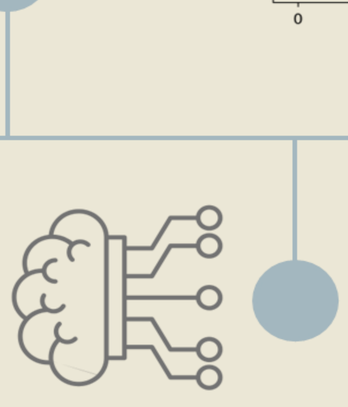
Semi-preparative biotransformations of 10 mmol substrate on 200 mL scale.

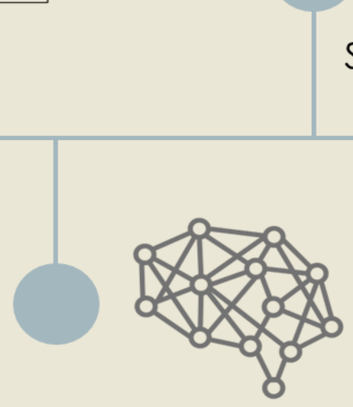
TARGET REACTION

CATALOPHORE

SCALE UP







COMPUTATIONAL ANALYSIS

Represenative panels cover substrates and IREDs with principal component analysis and multidimensional scaling for the substrate space, and sequence space analysis and model generation for the IRED space.

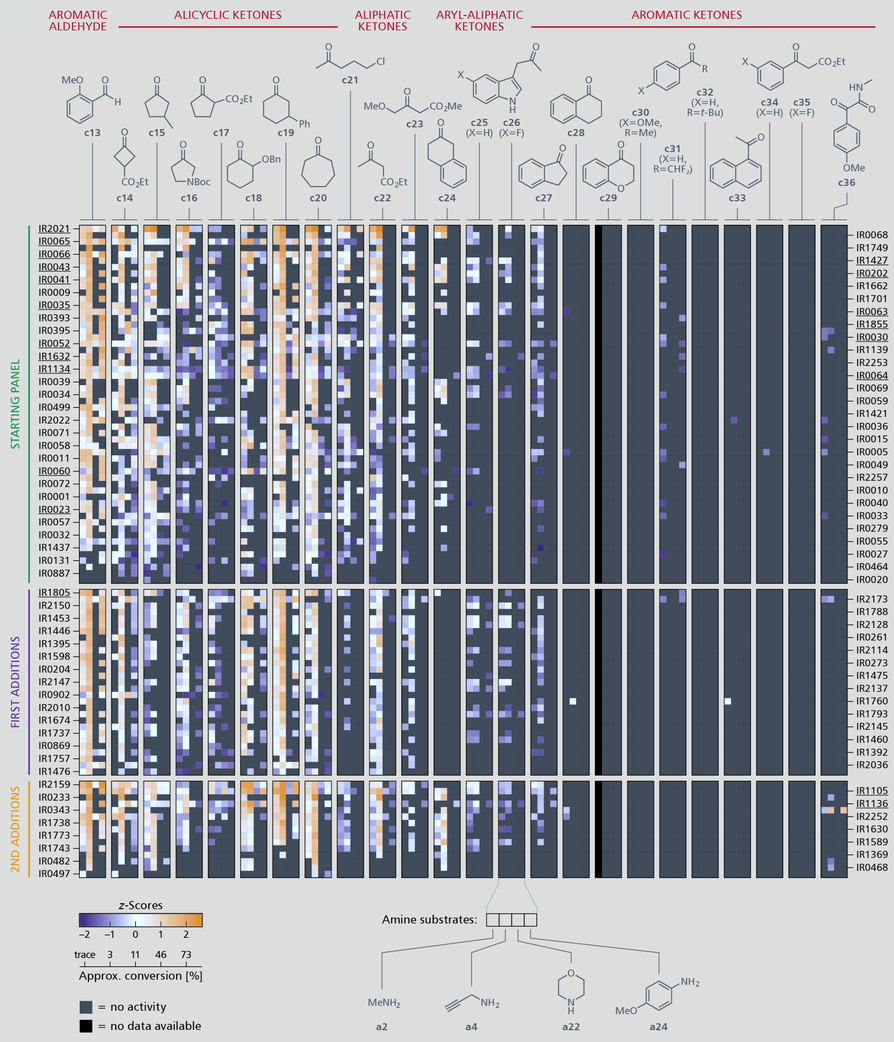
MACHINE LEARNING

ML-models correlated activities to differences in physico-chemical descriptors of the enzymes' active sites.

DEEP LEARNING

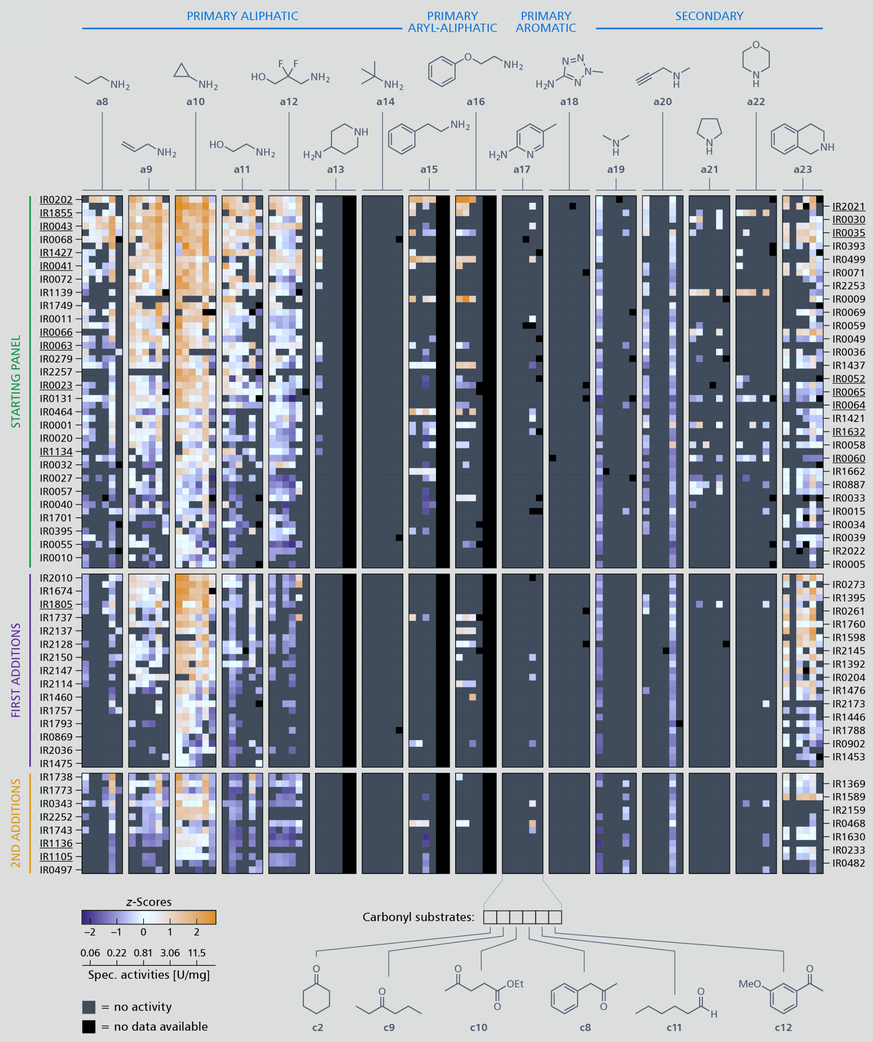
DL-embeddings and molecular features extracted using our catalophores were combined and fed into a downstream activity predictor.

EXPERIMENTAL VALIDATION



During the screening process[2], we delved into the reductive amination scope of the top 100 best-performing IREDs. This involved testing their capabilities with various substrates, including larger and more intricately functionalized compounds commonly employed as key building blocks in medicinal chemistry.

Strategy successfully guided IRED substrate scope expansion to reduce experimental screening efforts, and the models rationalized the observed and predicted activities of previously untested substrates.[3]



Heatmap representation of data from the screening of 100 enzymes against 24 carbonyls x 4 amines (carbonyl scope, **left**) and 6 carbonyl compounds x 16 amines (amine scope, **right**). Data are represented as z-scores (distance from mean in units of standard deviation) using a diverging purple-orange colour scale with white as the center point (mean, z = 0). Gray tiles represent inactive substrate-enzyme combinations, and black tiles indicate combinations for which no data are available.

References

1. Gilio, A. K. *et al*. Chem. Sci. 17, 4697 (2022); Casamajo, A. R. *et al*. JACS 145, 22041 (2023); Wetzl, D. *et al*. ChemCatChem 8, 2023 (2016); Velikogne, S. *et al.* ChemCatChem 16, 1749 (2016).
2. Berger, S. A. *et al.* ChemBioChem 24, e202300170 (2023).
3. Berger, S. A., Grimm, C. *et al.* manuscript in preparation.