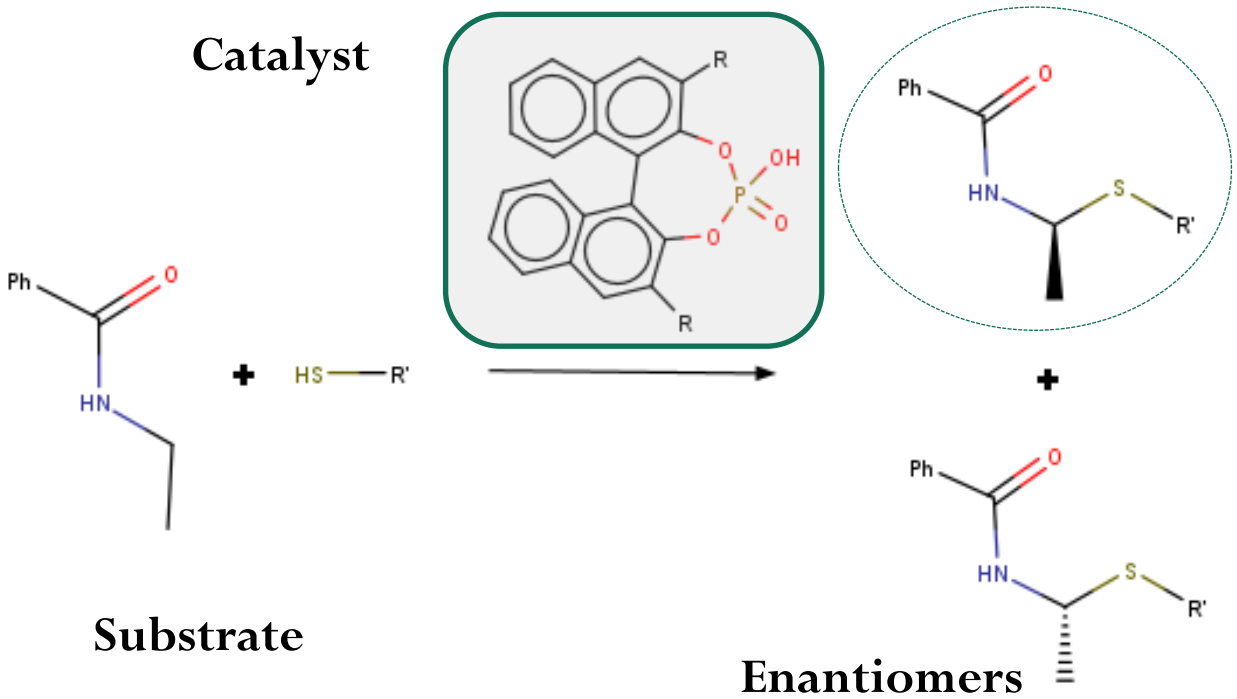
Beyond Black Boxes: Physics-Informed AI for Enantioselective Catalyst Design

Alexandre Varnek

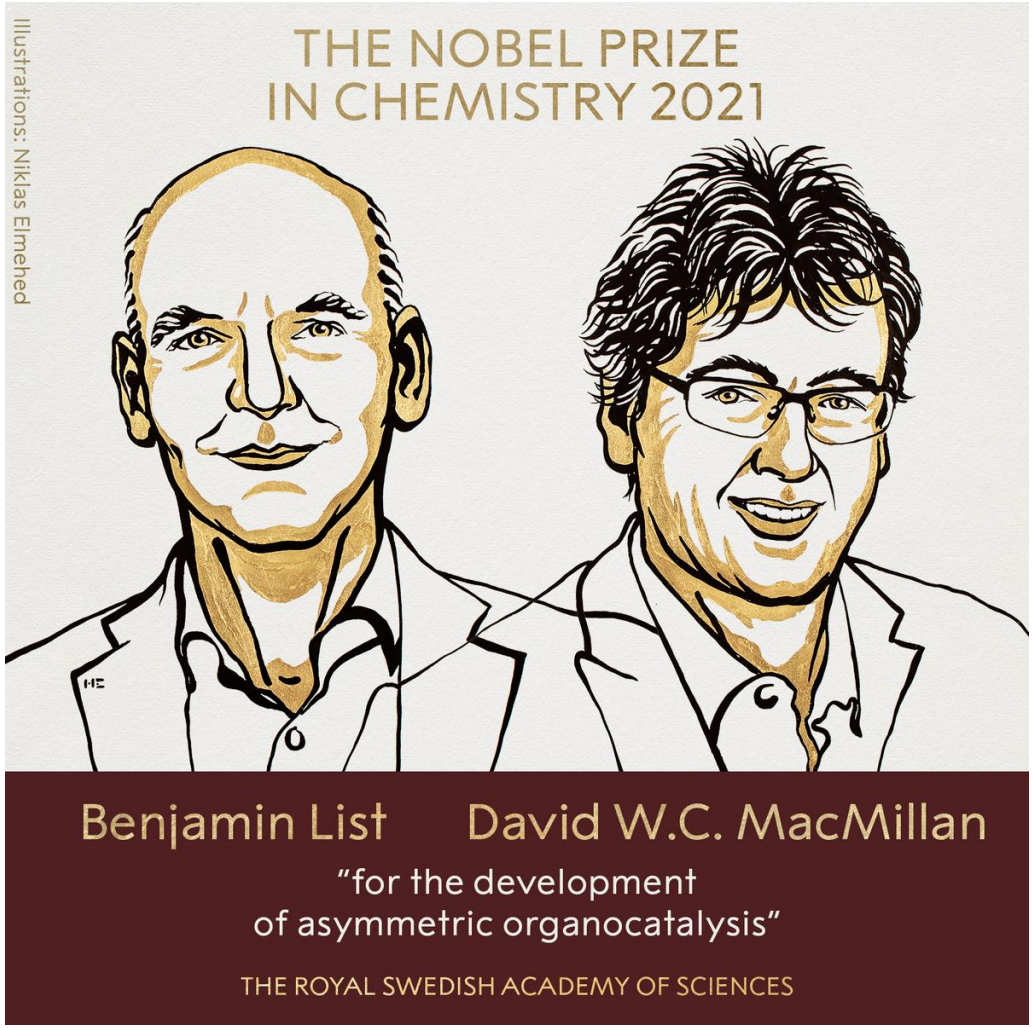
University of Strasbourg, France &
ICReDD, Hokkaido University, Japan

Design of new enantioselective catalysts



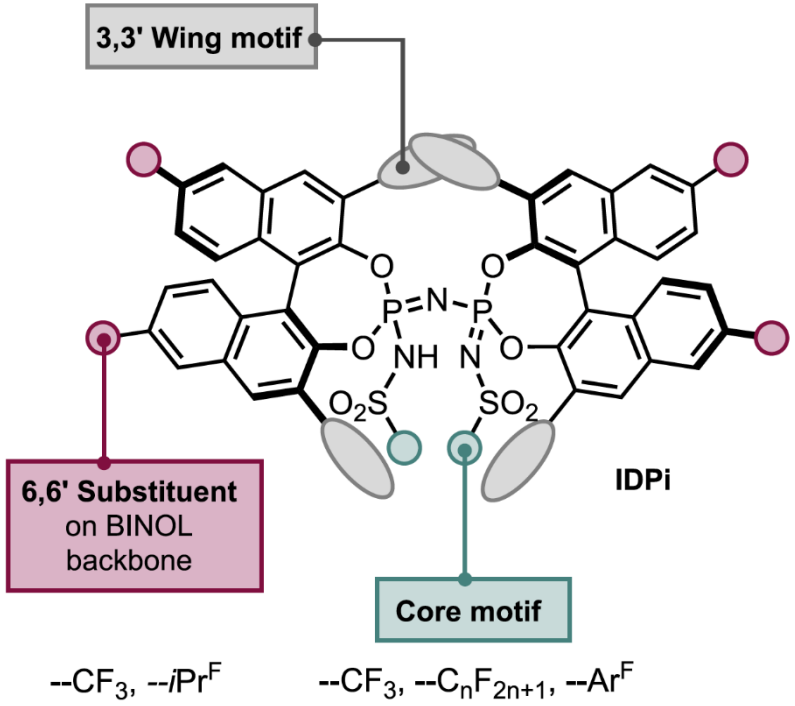
Measured property: enantiomeric ratio

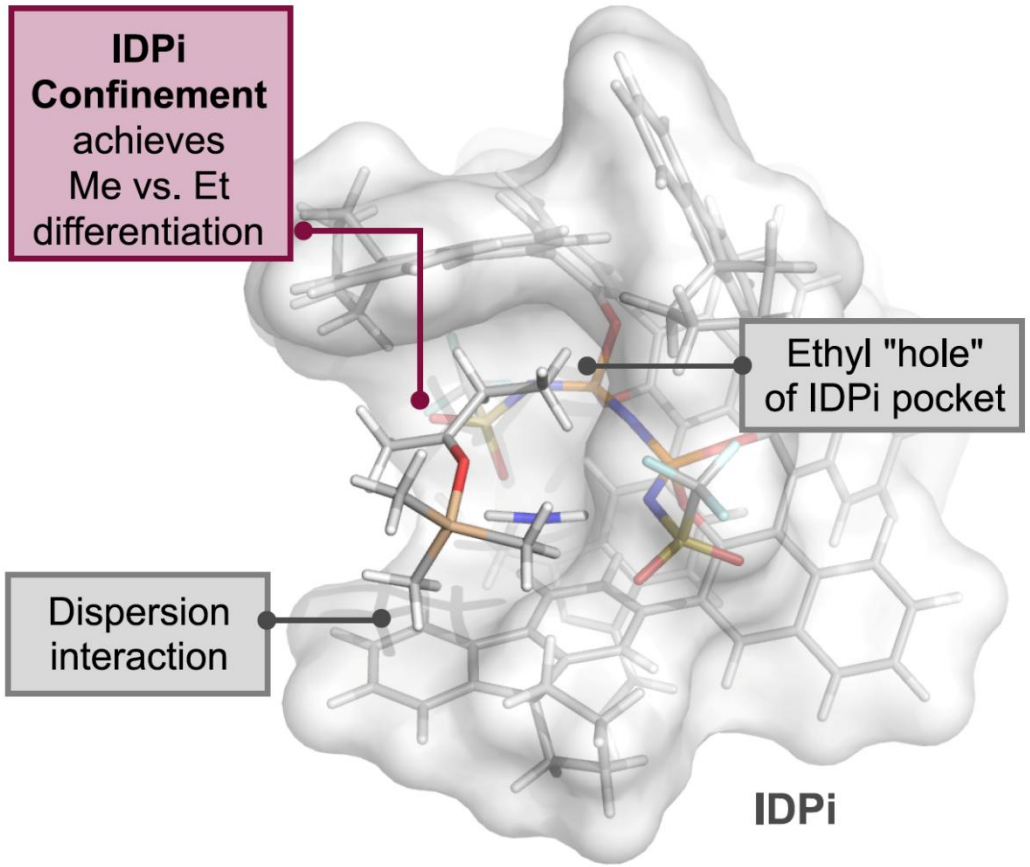
Modeled property: enantiomeric excess, $\Delta\Delta G$



Design of new enantioselective catalysts

IDPi (iminodiphosphorimidate) catalysts





Key features of IDPi catalysts:

High acidity, Chiral confinement , Tunability

P. Kaib et al., *Angew. Chem. Int. Ed.* 2016, **55**, 13200–13203

Data-Driven and Hypothesis-Driven paradigms

$$i\hbar\frac{\partial}{\partial t}|\Psi\rangle = \hat{H}|\Psi\rangle$$

Top-Down
Theory-driven approach

Inference

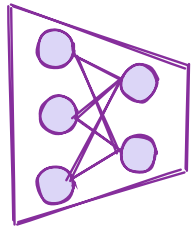
Theory

Physics-informed Machine Learning

Data

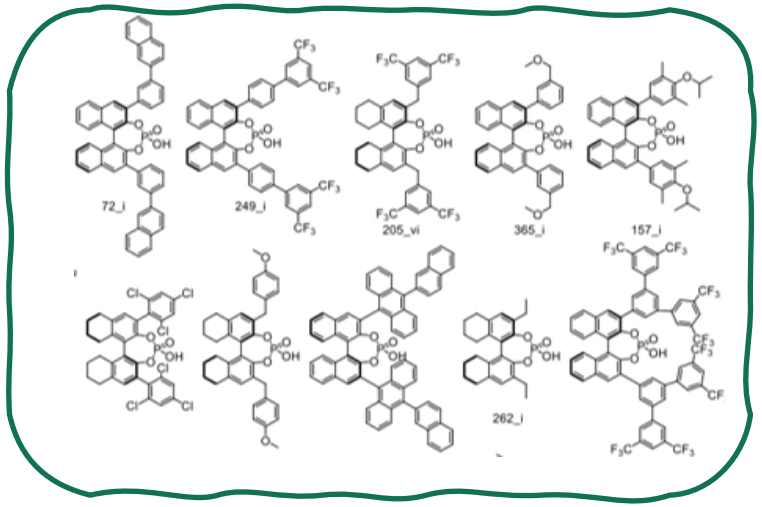
Inference

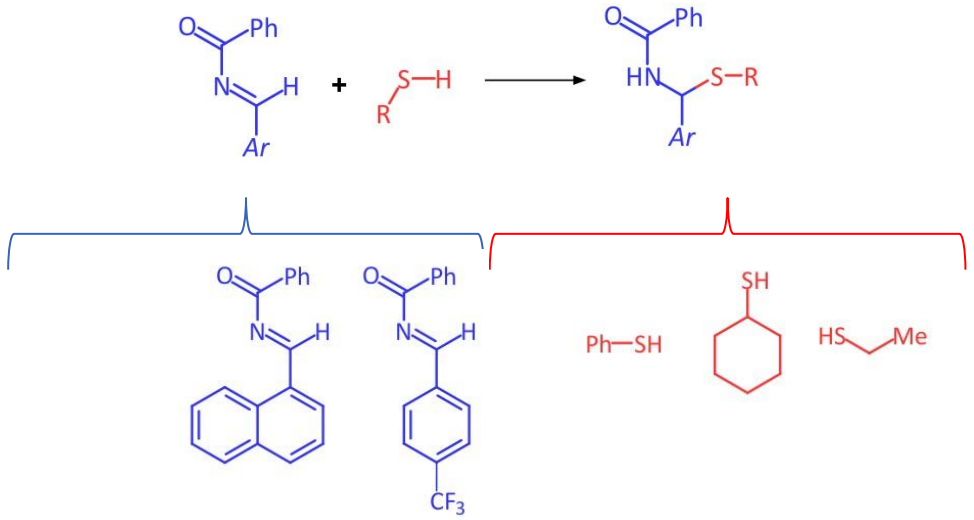
Bottom-Up
Data-driven approach



Design of enantioselective catalysts: methodology

N catalysts are used in M reactions (substrates) forming N*M catalyst-reaction pairs

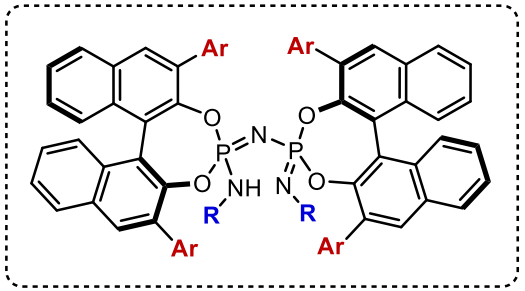


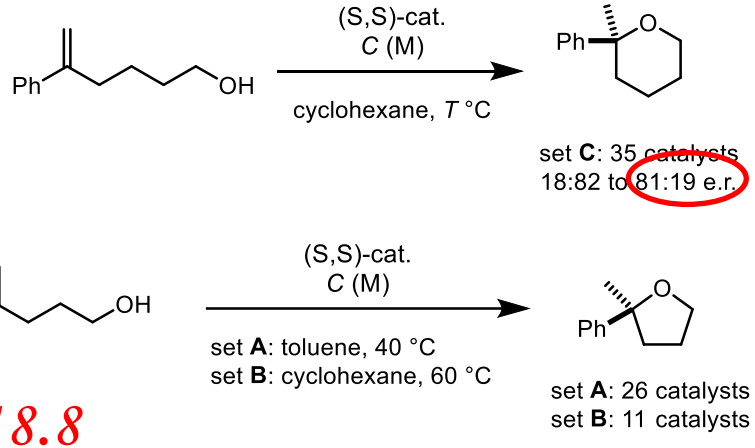


Descriptor vector for a *catalyst-reaction-condition* combination

Reaction	*Catalyst*	*Solvent*	*Temperature*
2D descriptors from CGR	2D of 3D descriptors	exp. parameters	T, K

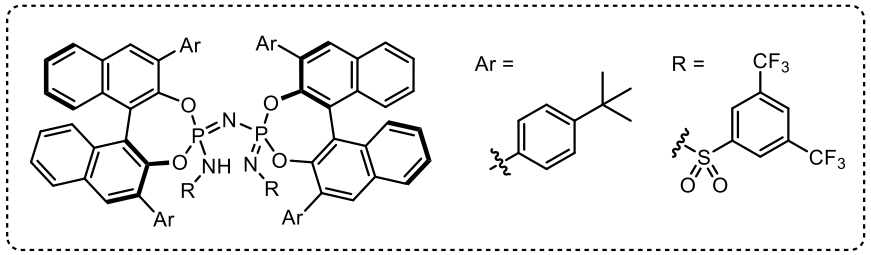
Design of enantioselective catalysts for asymmetric tetrahydropyran synthesis

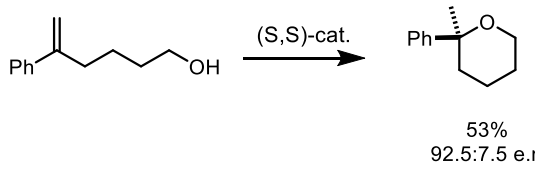




best catalyst: *ER* = *81.2:18.8*

- SVR model based on tunable ISIDA/CGR descriptors
- Screening of small virtual combinatorial library
- Experimental tests





In silico* designed catalyst: *ER* = *92.5:7.5

ER* – *enantioselectivity ratio

N. Tsiji *et al. Angew. Chemie* 2023, 62 (11), e202218659

Multi-Instance Learning (MIL)

DOI: 10.1002/wcms.1698

ADVANCED REVIEW

WIREs COMPUTATIONAL MOLECULAR SCIENCE WILEY

Chemical complexity challenge: Is multi-instance machine learning a solution?

Dmitry Zankov[1] | Timur Madzhidov[2] | Alexandre Varnek[1,3] | Pavel Polishchuk[4]

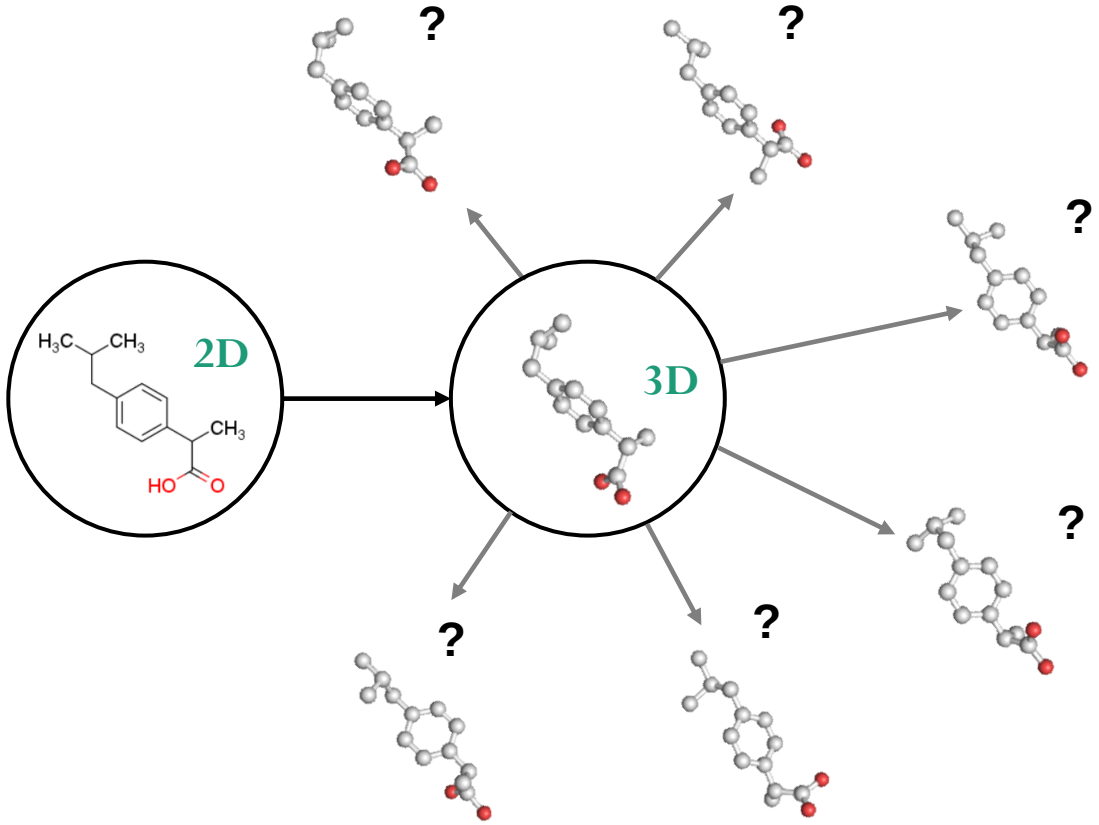
JCIM JOURNAL OF CHEMICAL INFORMATION AND MODELING

pubs.acs.org/jcim Article

QSAR Modeling Based on Conformation Ensembles Using a Multi-Instance Learning Approach

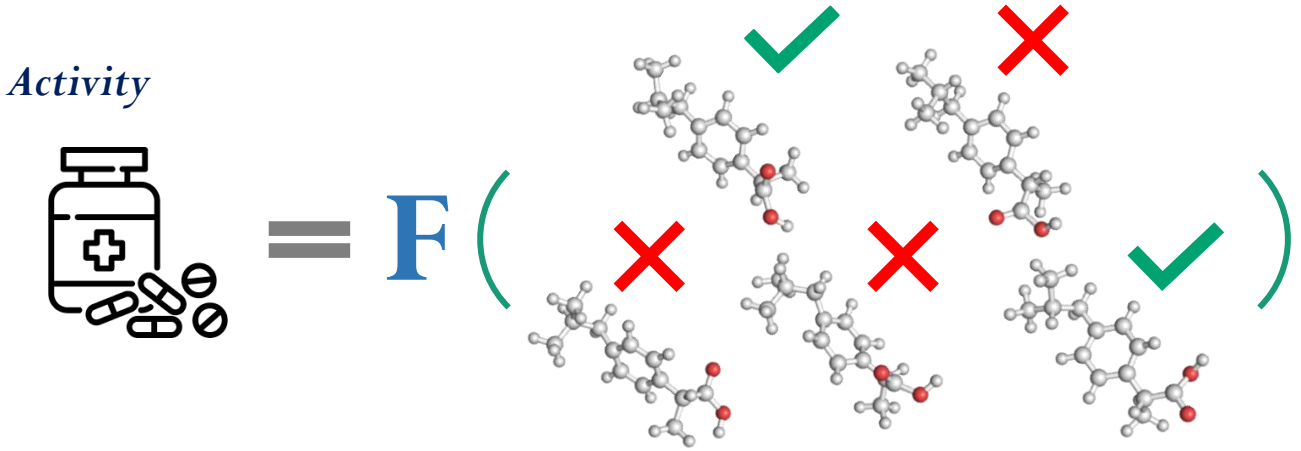
Dmitry V. Zankov, Mariia Matveieva, Aleksandra V. Nikonenko, Ramil I. Nugmanov, Igor I. Baskin, Alexandre Varnek,* Pavel Polishchuk,* and Timur I. Madzhidov*

Cite This: *J. Chem. Inf. Model.* 2021, 61, 4913–4923 Read Online

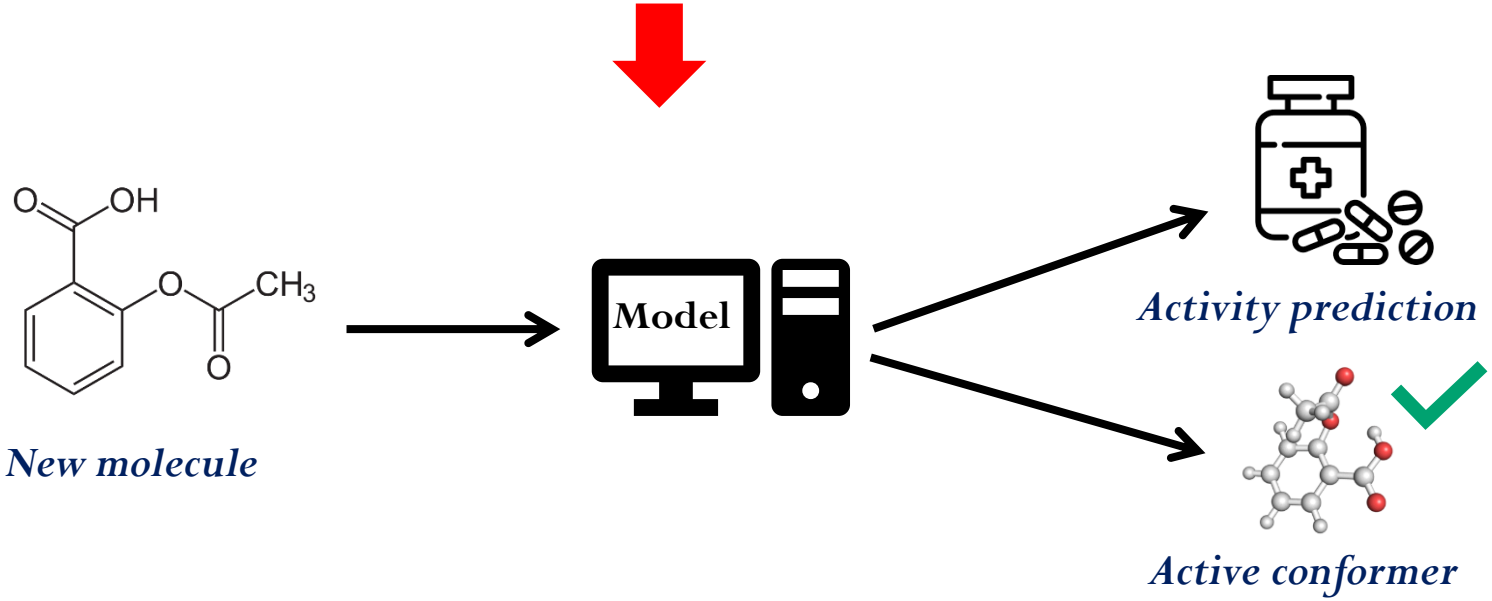


MIL considers ensemble of low energy conformers

Multi-Instance Learning

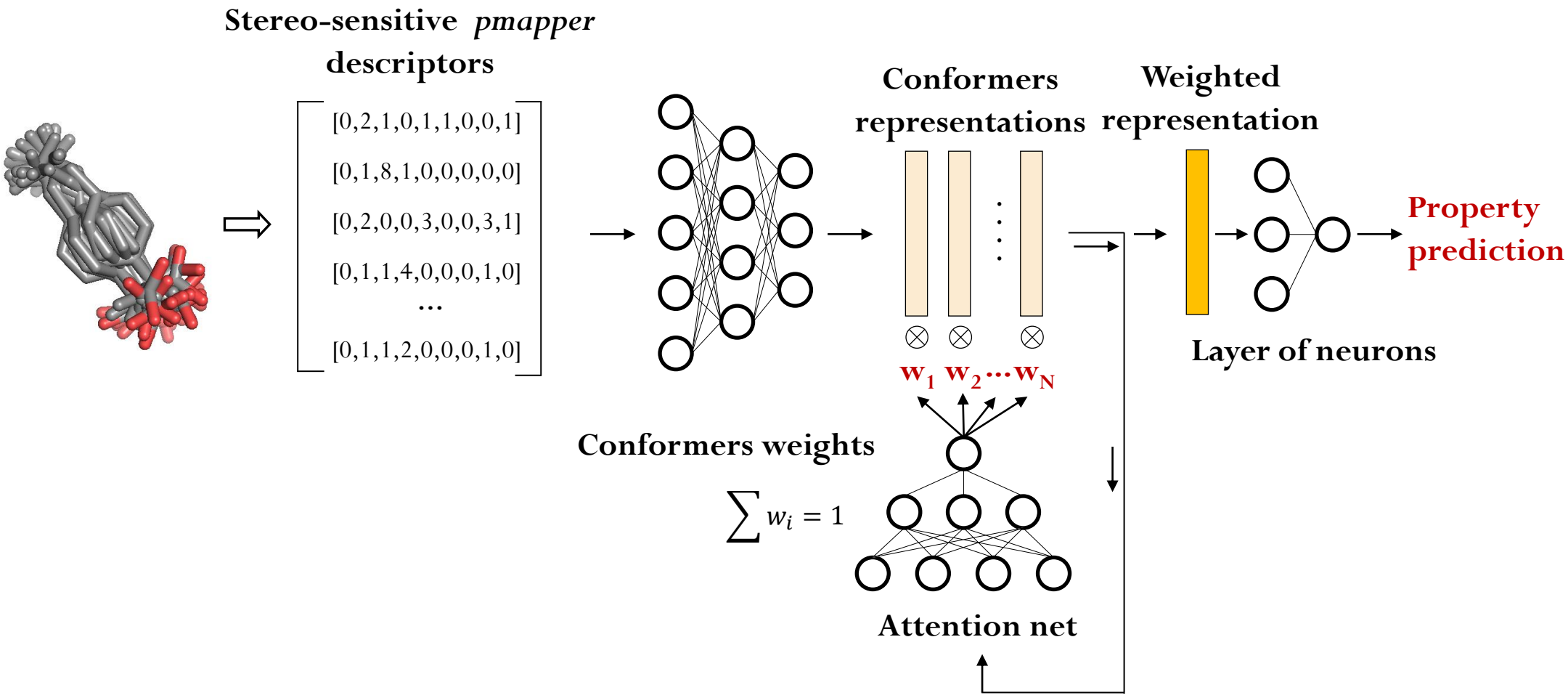


- *No conformers selection*
- *No conformers alignment*
- *Key conformers identification*



New molecule

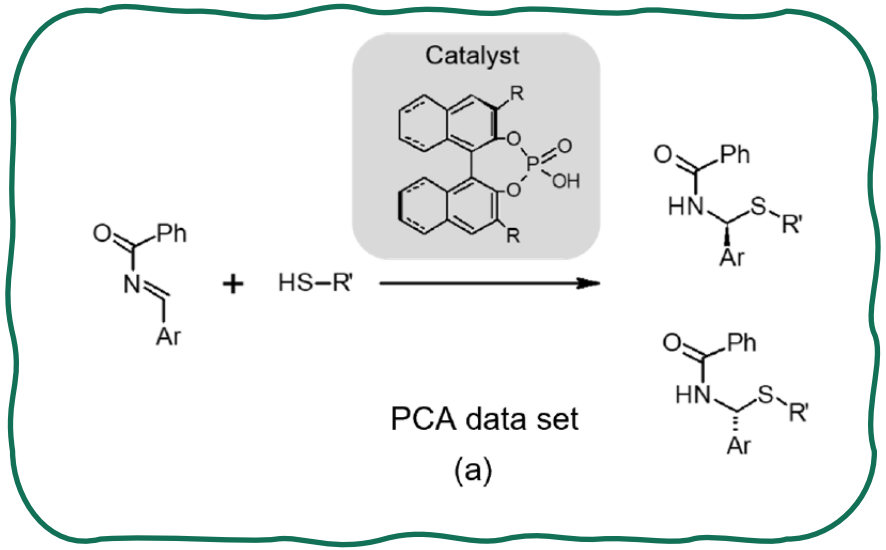
Attention-based multi-instance neural network

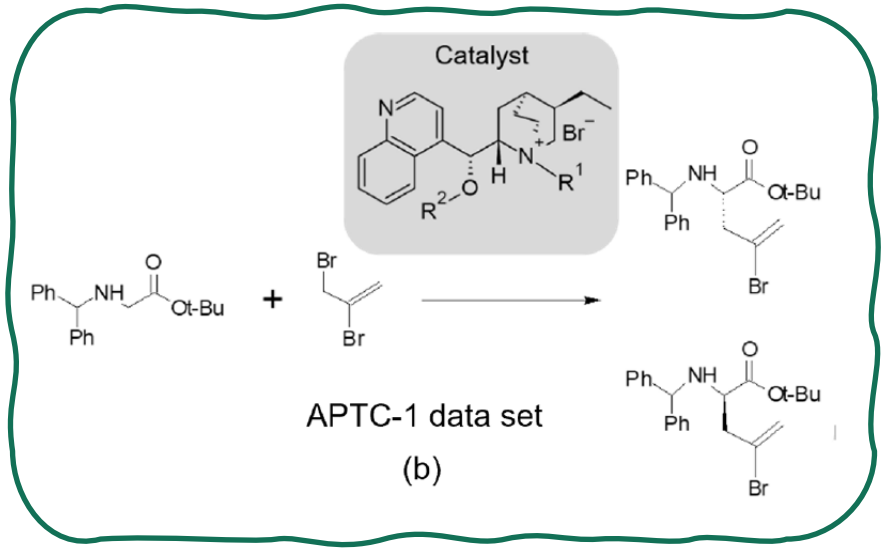


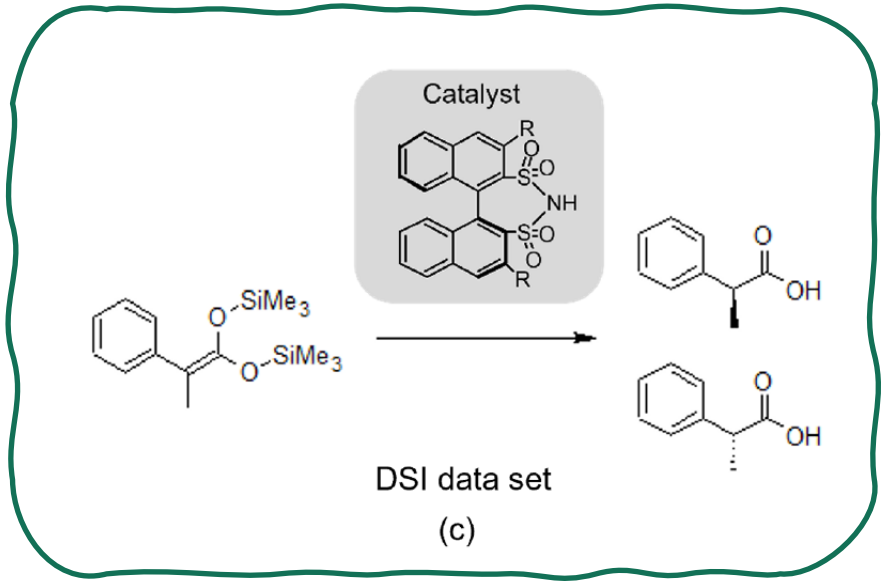
D. Zankov et al J. Chem. Inf. Model. 2021, 61, 4913−4923

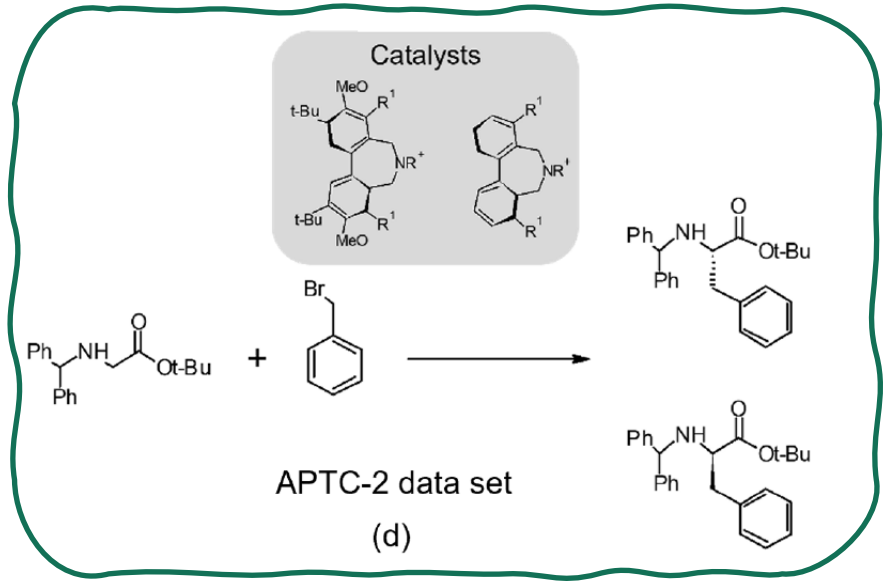
Catalysts: MIL models of enantioselectivity

Dmitri ZANKOV



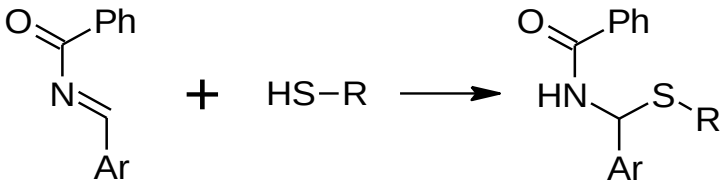




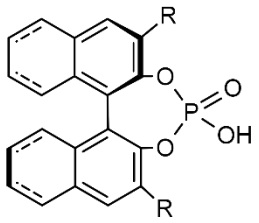


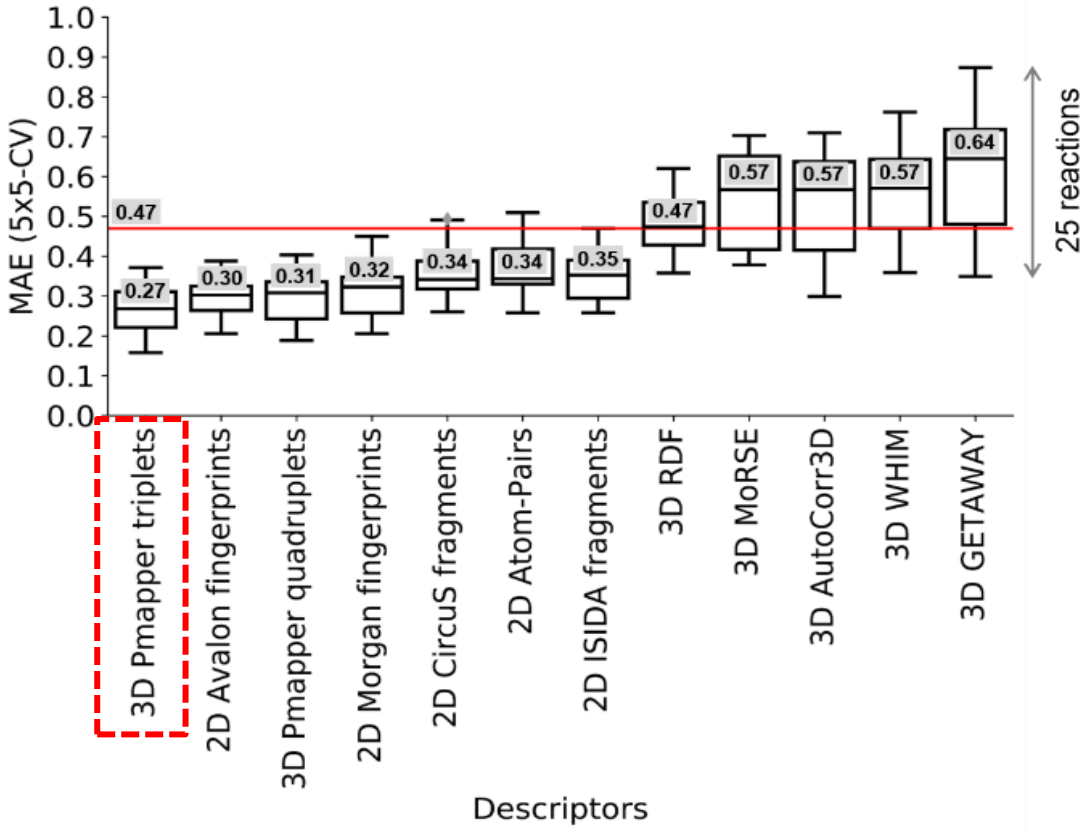
D. Zankov, et al (2023). J. Chem. Inf. Model., https://doi.org/10.1021/acs.jcim.3c00393

Benchmarking study on literature data



Catalyst

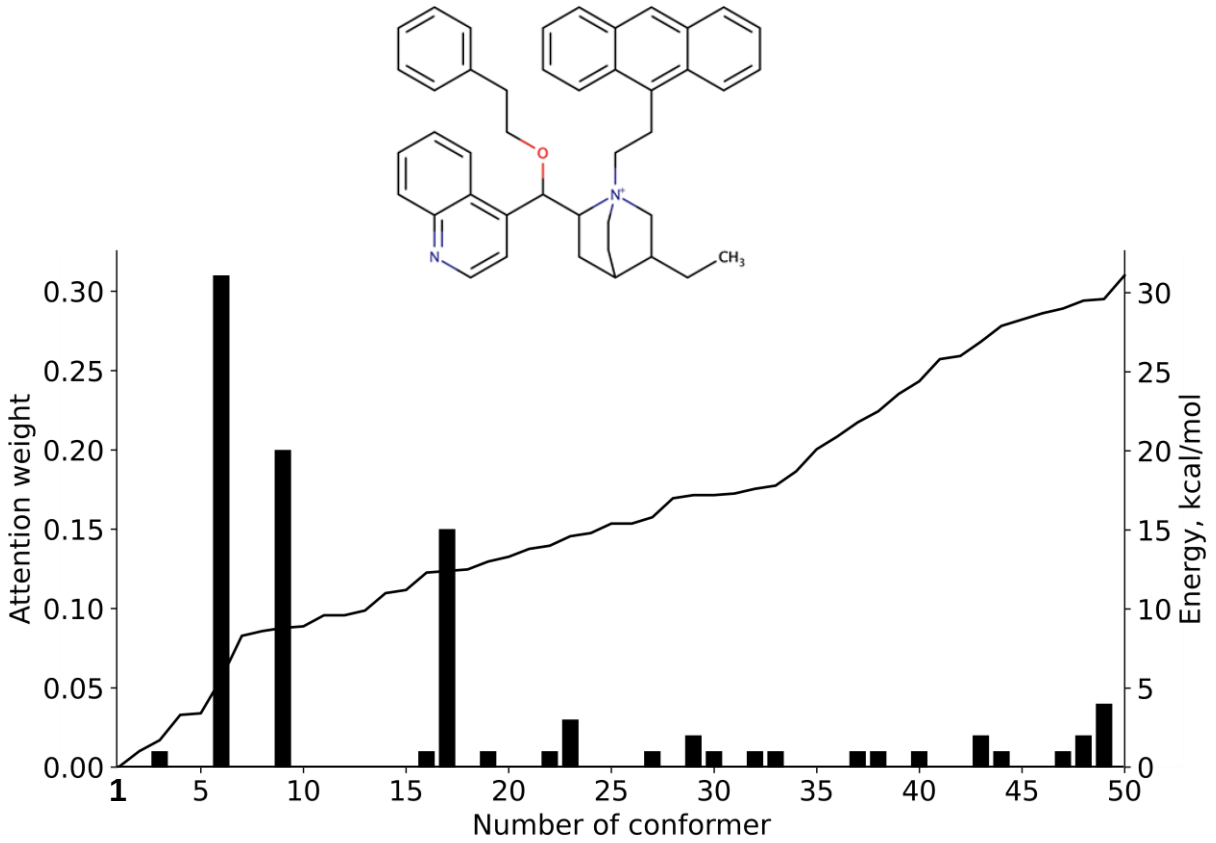




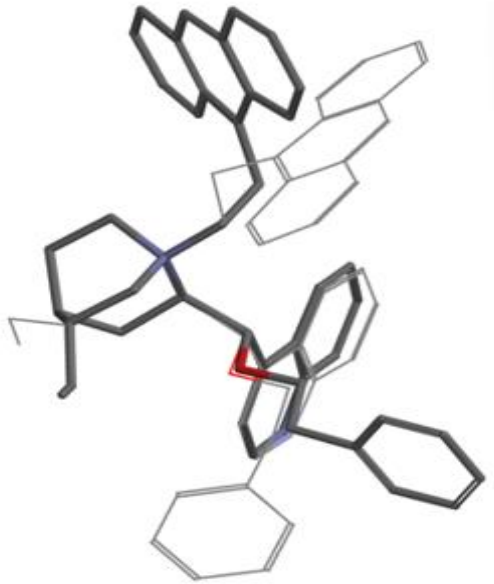
MIL /3D models outperform single-instance models on 2D or 3D descriptors

D. Zankov, et al (2023). *J. Chem. Inf. Model*., https://doi.org/10.1021/acs.jcim.3c00393

Catalysts: Key instances identification



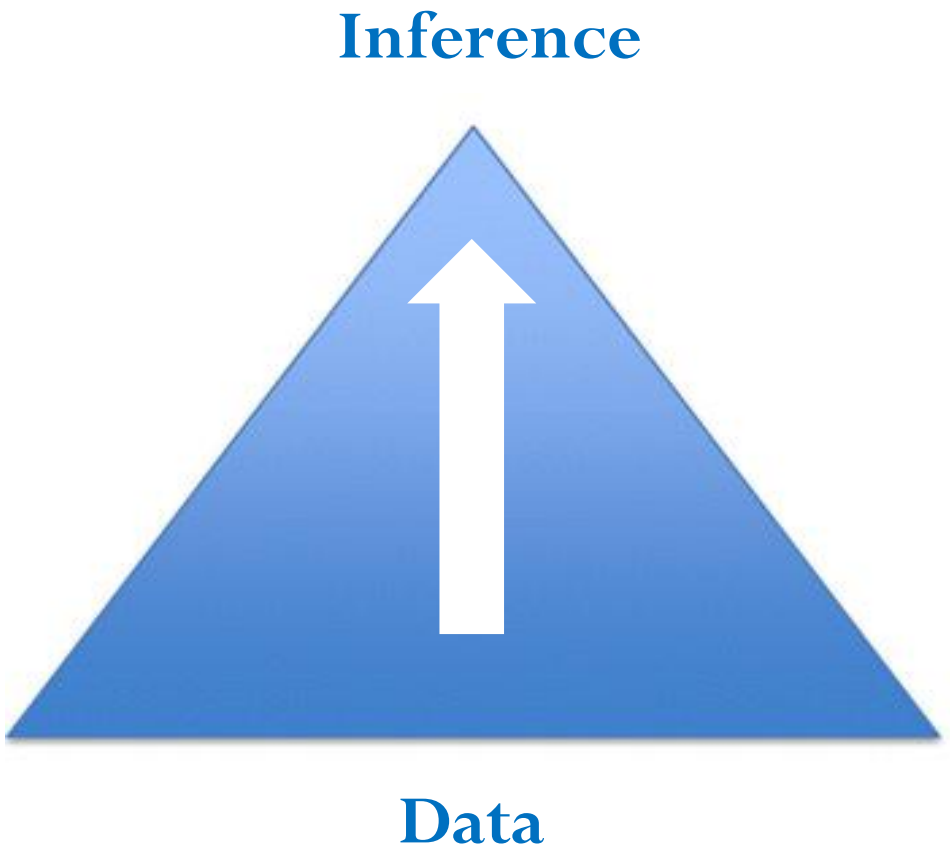
"Importance" weights assigned to the conformers and corresponding relative energies



Aligned conformer with the largest weight (**in bold**) and the lowest energy conformer.

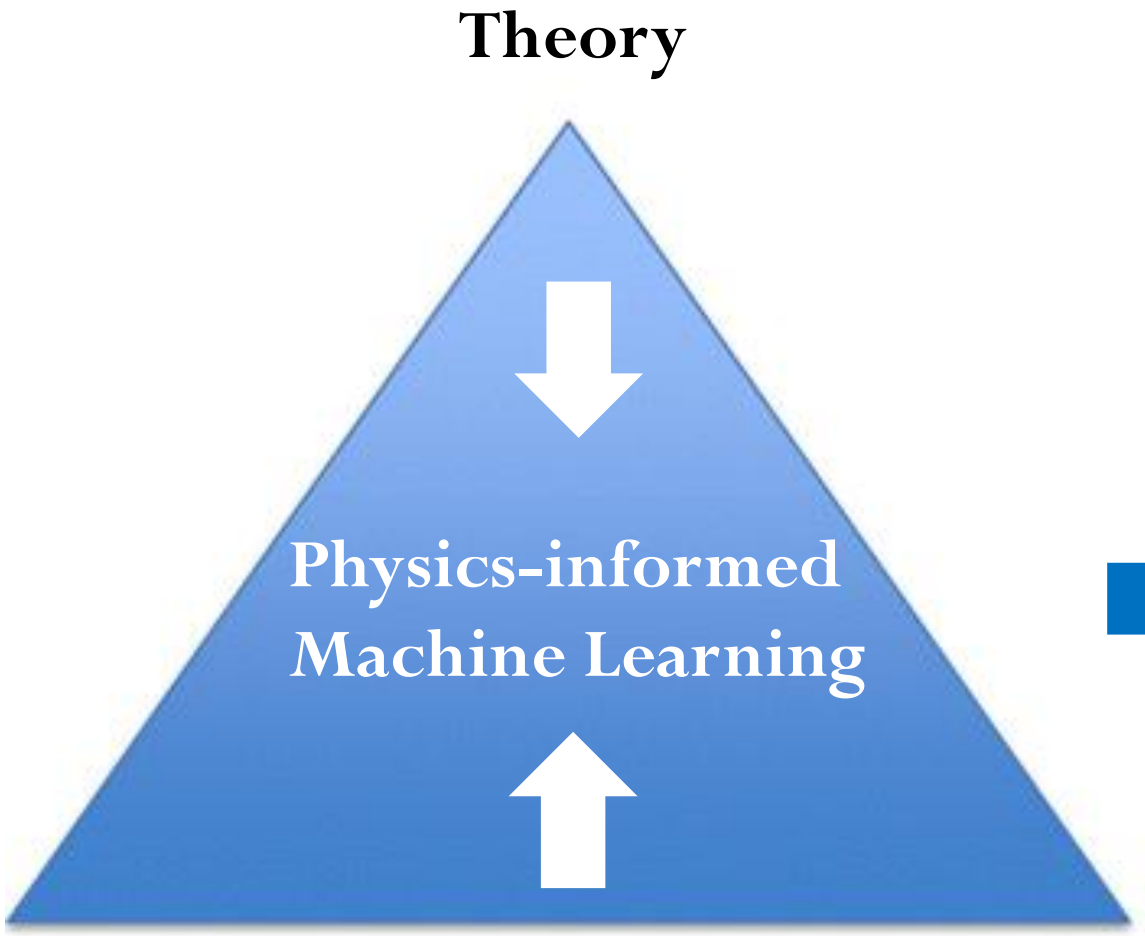
D. Zankov, et al (2023). *J. Chem. Inf. Model.*, https://doi.org/10.1021/acs.jcim.3c00393

Limitation of Data-Driven approach



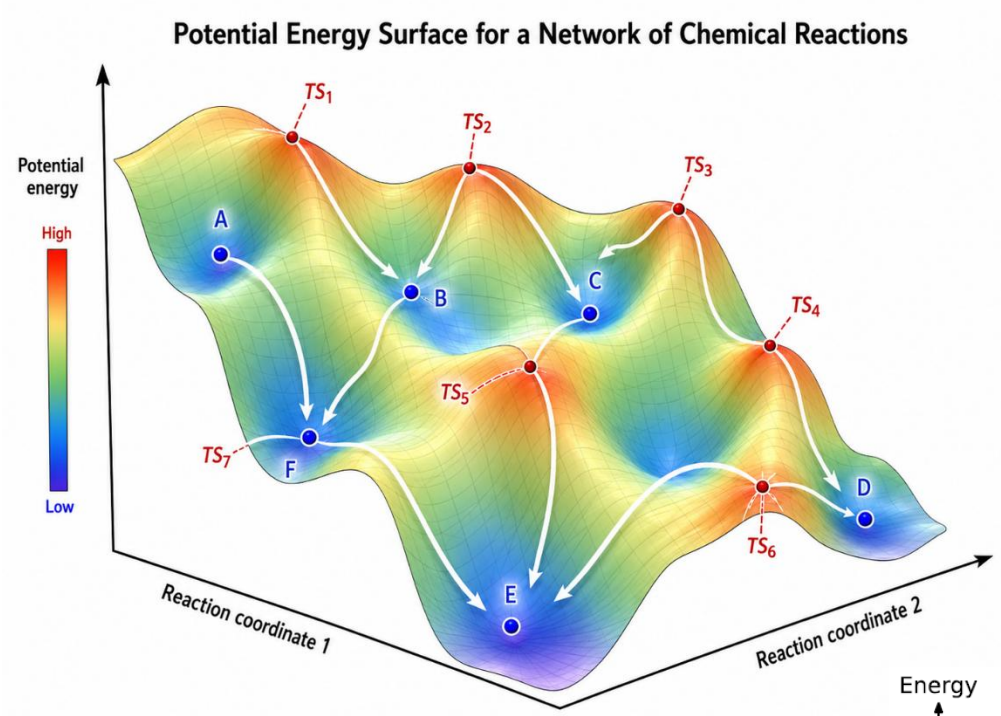
- Lack of experimental data
- Poor interpretability

Combining Data- and Hypothesis-Driven approaches

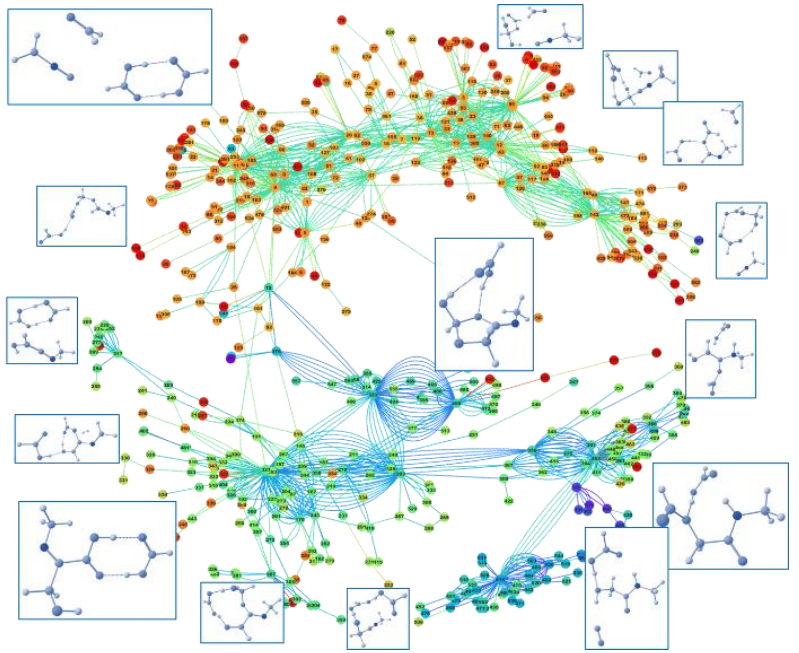


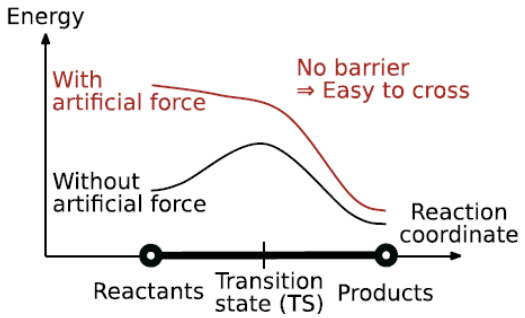
ML models on QM data

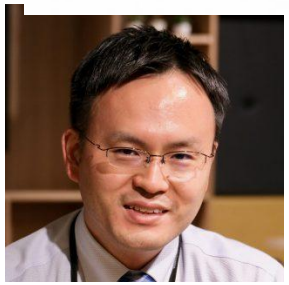
Artificial Force Induced Reaction (AFIR)



Reaction path network





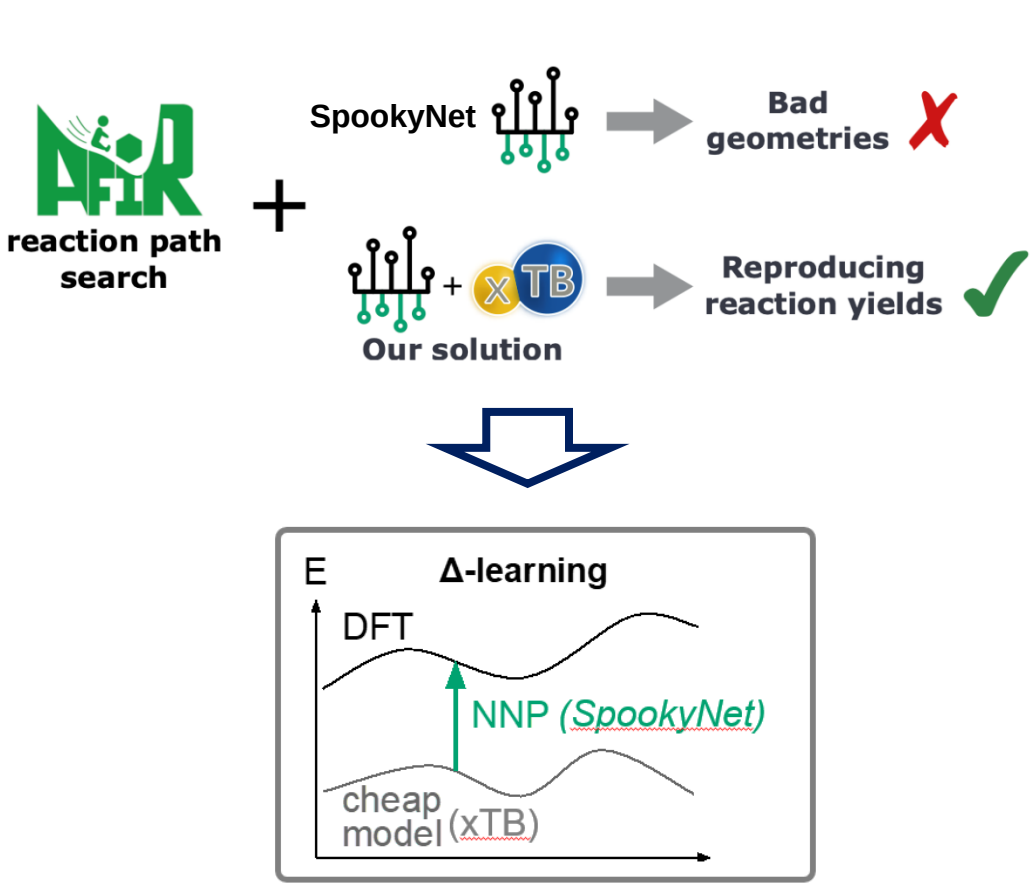


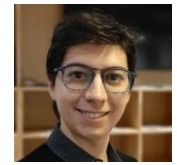
Satoshi MAEDA

Maeda, S.; Harabuchi, Y. *WIREs Comput. Mol. Sci.*, 2021, *11*, e1538.

Courtesy of Prof. S. Maeda

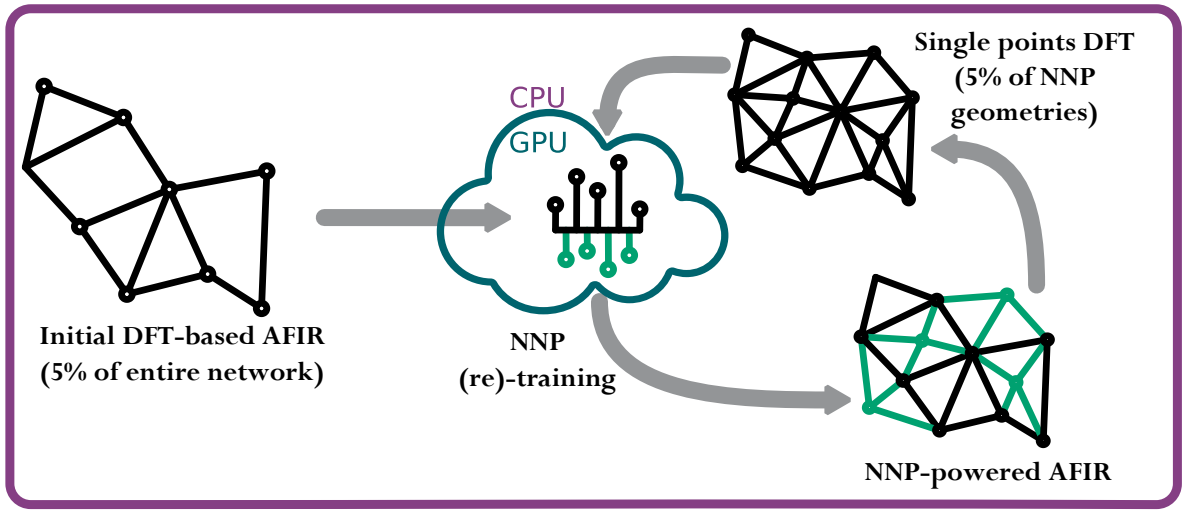
Designing robust potentials for AFIR search acceleration





Ruben Staub

Iterative reaction network exploration



- *Accelerate AFIR-based exploration*
  ⇒ ≈ 10^3 speedup (1 month → few hours)
- *Expand exploration capabilities*
  ⇒ Increase system size (≈ 20 → 250 atoms)

Unke, O.T. et al. *Nat. Commun.* 2021, 12, 7273

Staub, R.; Gantzer, P.; Harabuchi, Y.; Maeda, S.; Varnek, A., *Molecules*, 2023.

Predicting IDPi Enantioselectivity via Gigantic Reaction Path Networks

ACS central science

Open Access

This article is licensed under CC-BY 4.0

http://pubs.acs.org/journal/acscii

Article

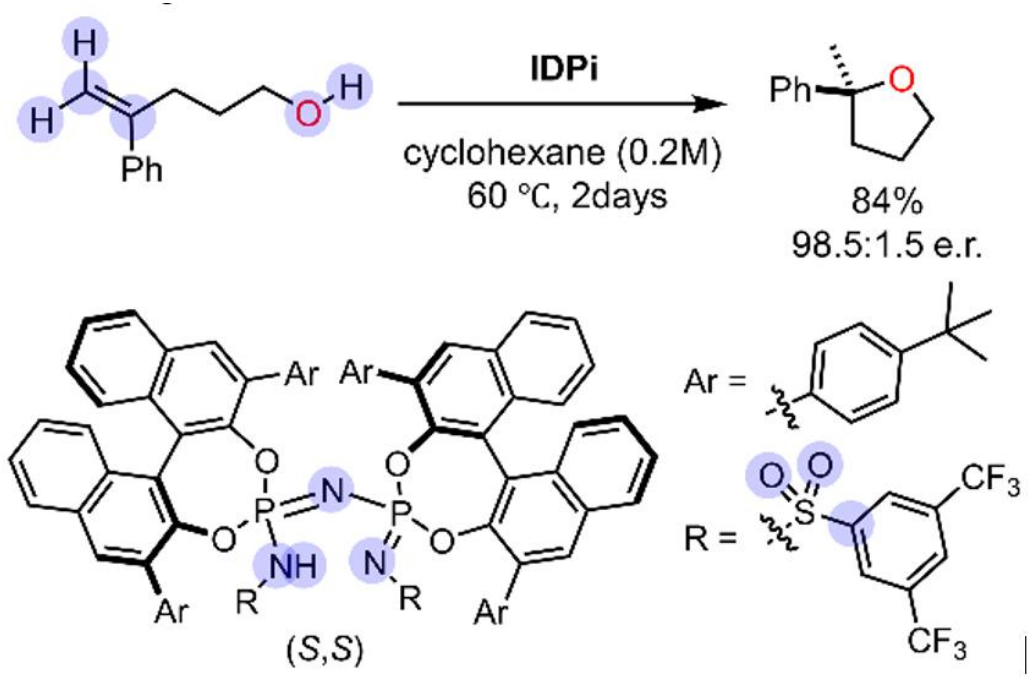
Predicting Enantioselectivity via Kinetic Simulations on Gigantic Reaction Path Networks

Yu Harabuchi, Ruben Staub, Min Gao, Nobuya Tsuji, Benjamin List, Alexandre Varnek, and Satoshi Maeda*

Cite This: *ACS Cent. Sci.* 2026, 12, 524–531

Read Online

Asymmetric hydroalkoxylation reaction



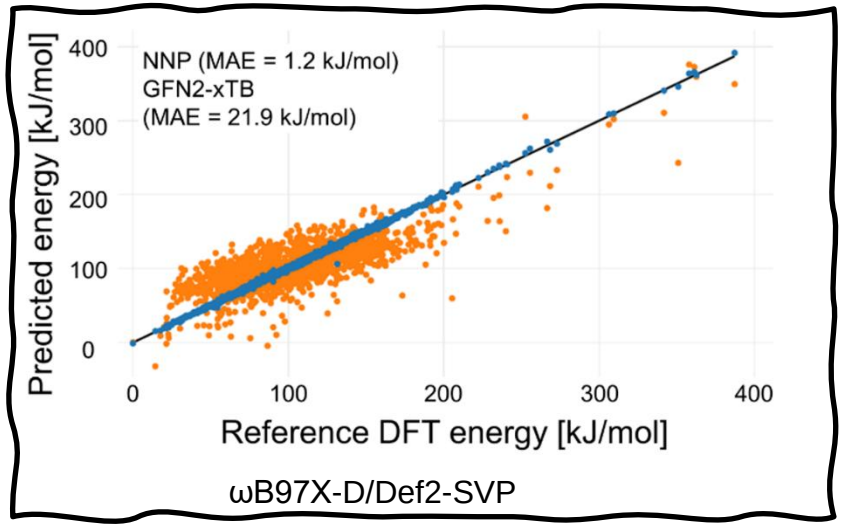
NNP modeling

228 atoms

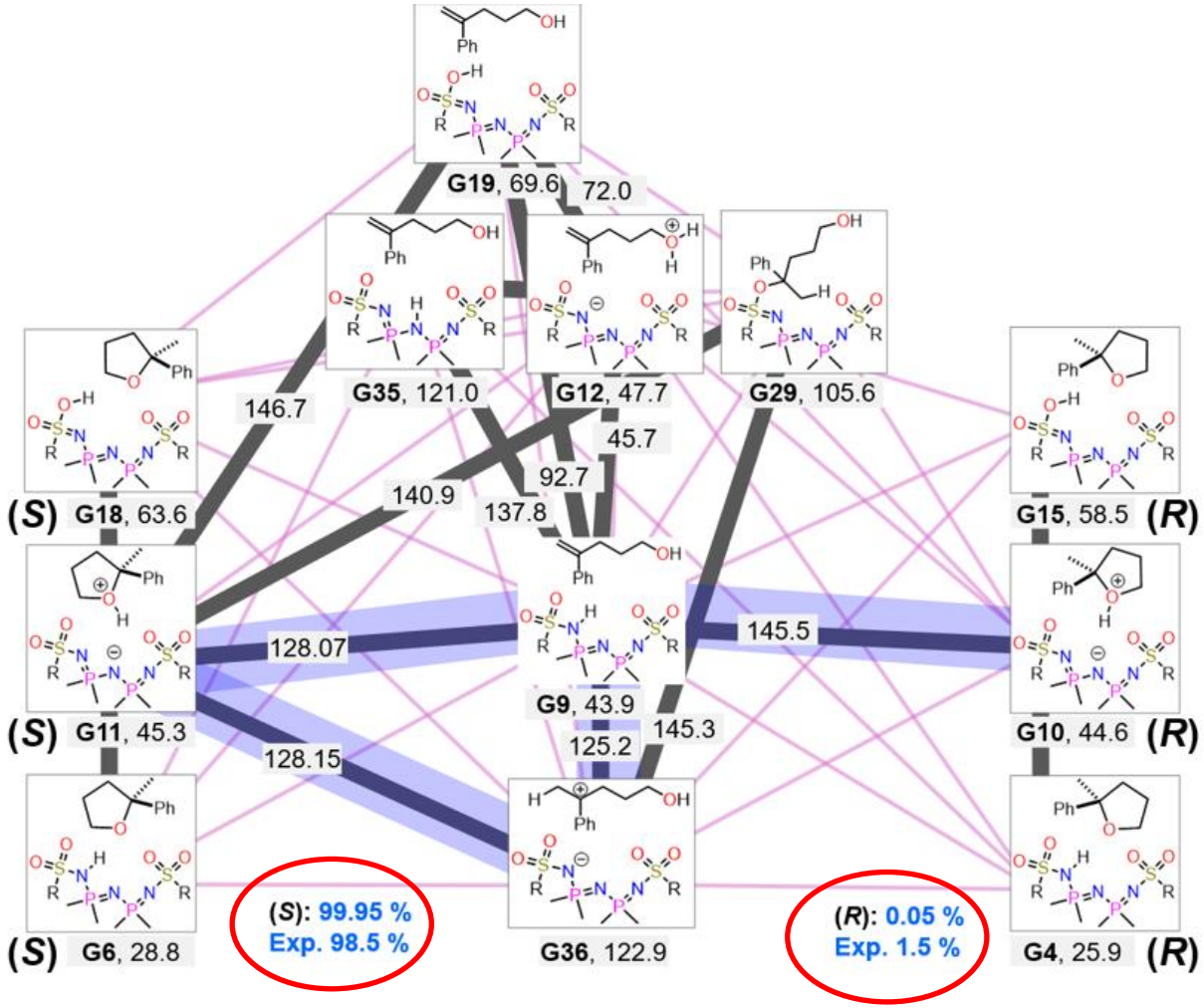
20 296 Equilibrium states

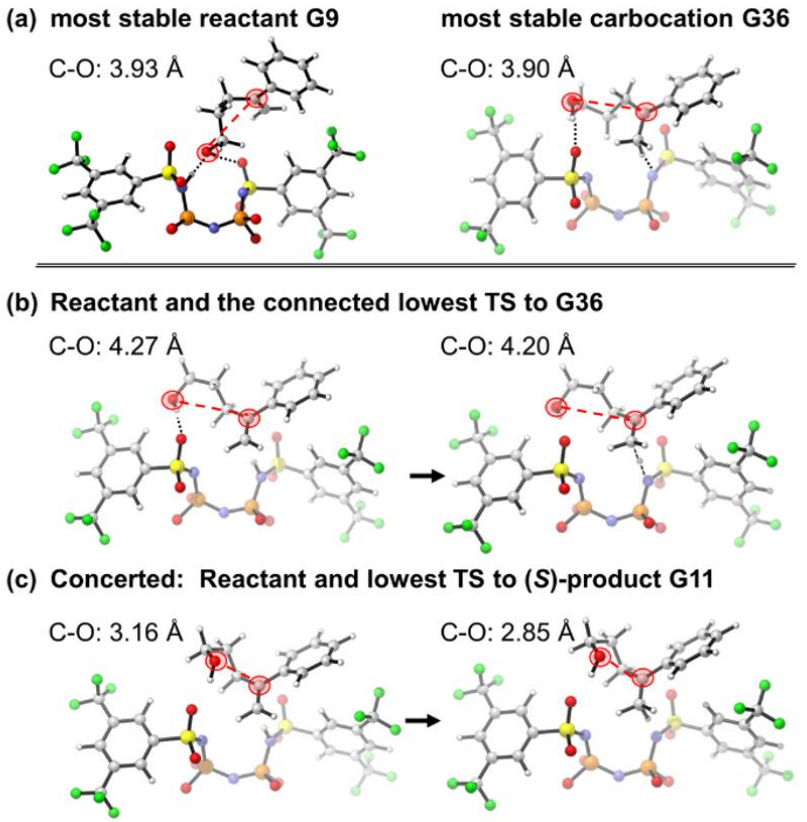
48 552 Transition states

48 463 paths



Predicting Enantioselectivity via Gigantic Reaction Path Networks

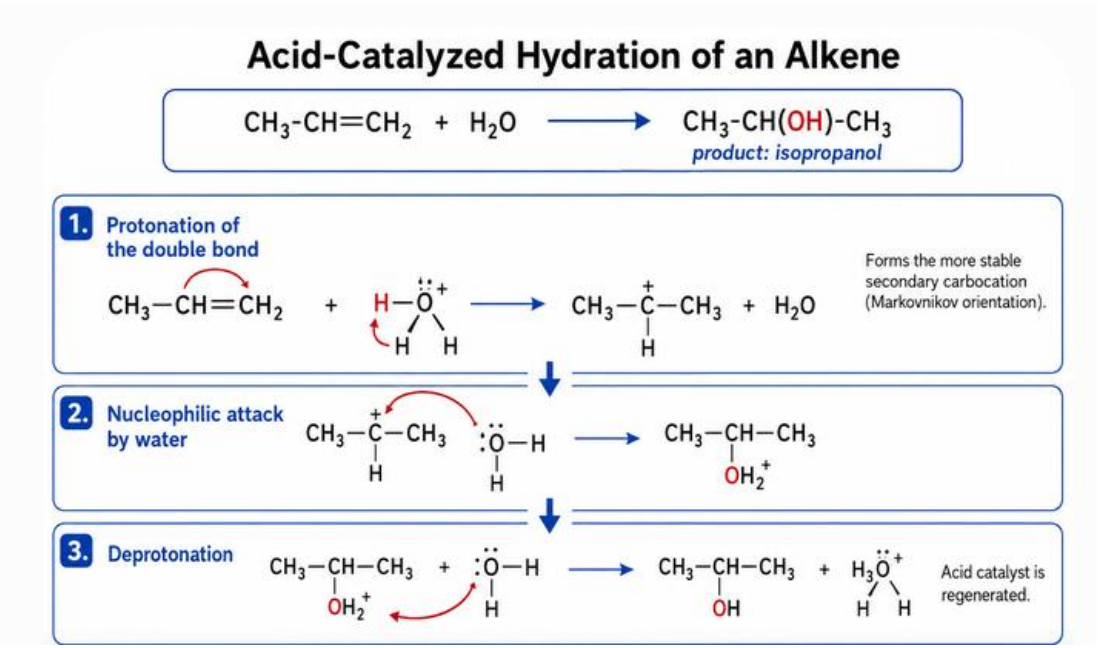




ΔΔG (R/S) = 17.4 kJ/mol (calc)
11.6 kJ/mol (exp)

Yu Harabuchi et al. *ACS Centr. Sci.*. 2026, **12**, 524

Catalytic Asymmetric Hydration of Alkenes

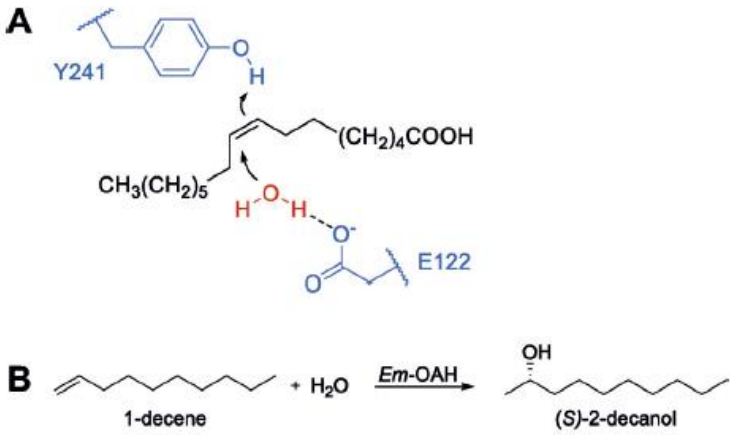


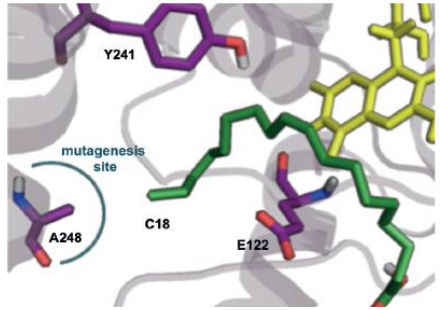
Enantioselective Reactions

International Edition: DOI: 10.1002/anie.201810005
German Edition: DOI: 10.1002/ange.201810005

Asymmetric Enzymatic Hydration of Unactivated, Aliphatic Alkenes

*Rebecca M. Demming, Stephan C. Hammer, Bettina M. Nestl, Sebastian Gergel, Silvia Fademrecht, Jürgen Pleiss, and Bernhard Hauer**





oleic acid hydratase

H. Hatano et al JACS, 2026, DOI: 10.1021/jacs.6c06916

Catalytic Asymmetric Hydration of Alkenes

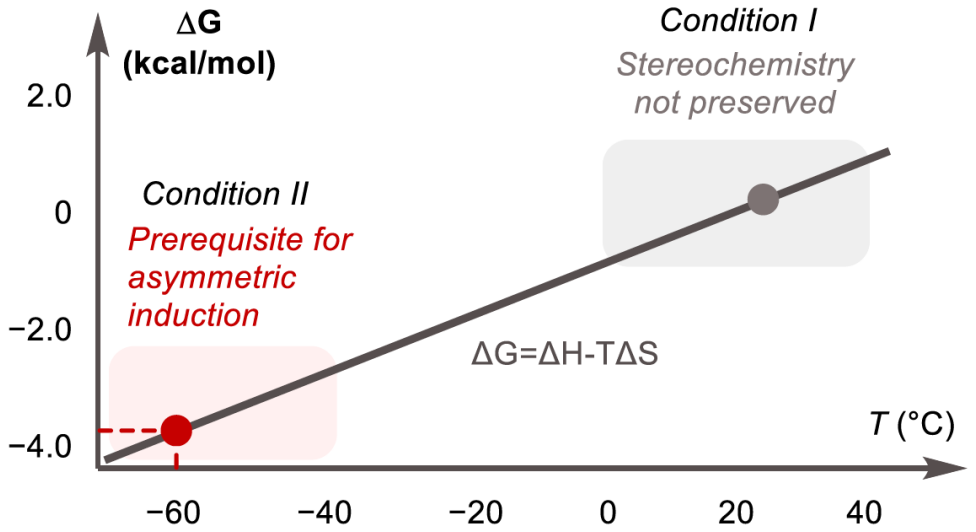
a **Asymmetric Markovnikov Hydration and Thermodynamic Considerations**

Olefins
abundant feedstocks

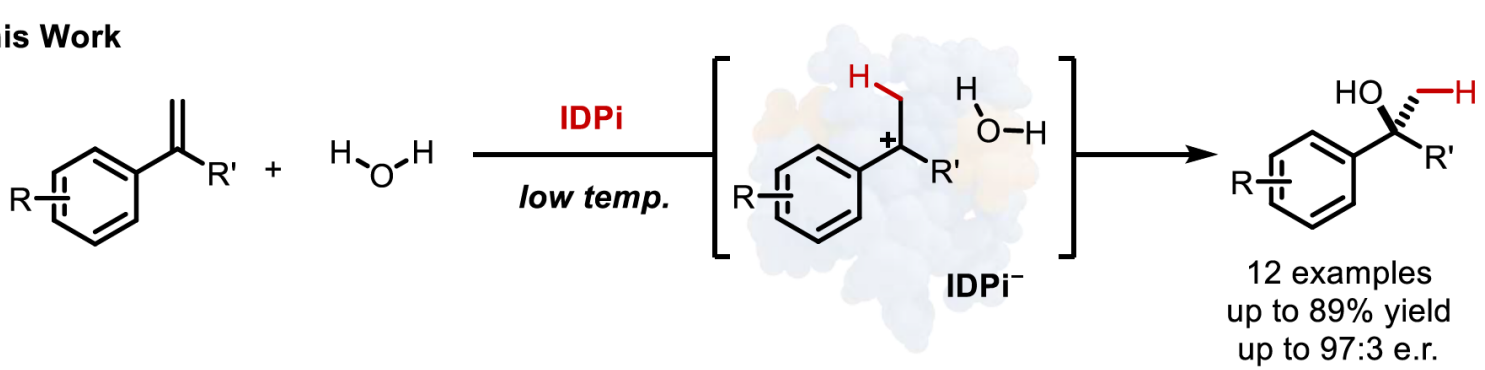
Shifting of equilibrium
via cryogenic conditions

Alcohols
enantioenriched

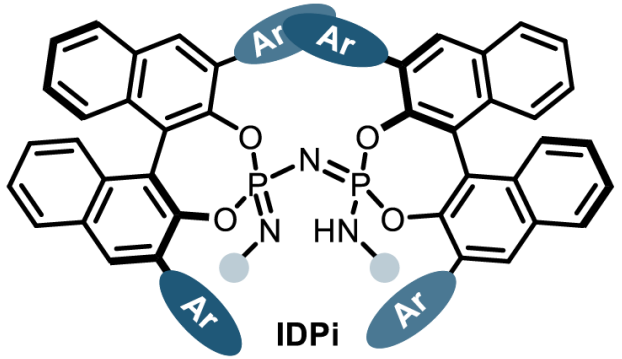
T	ΔG (kcal/mol)	for hydration of α-butylstyrene
25 °C	+0.1	*rapid equilibration*
−60 °C	***−3.2***	***Thermodynamic bias toward product***



b **This Work**

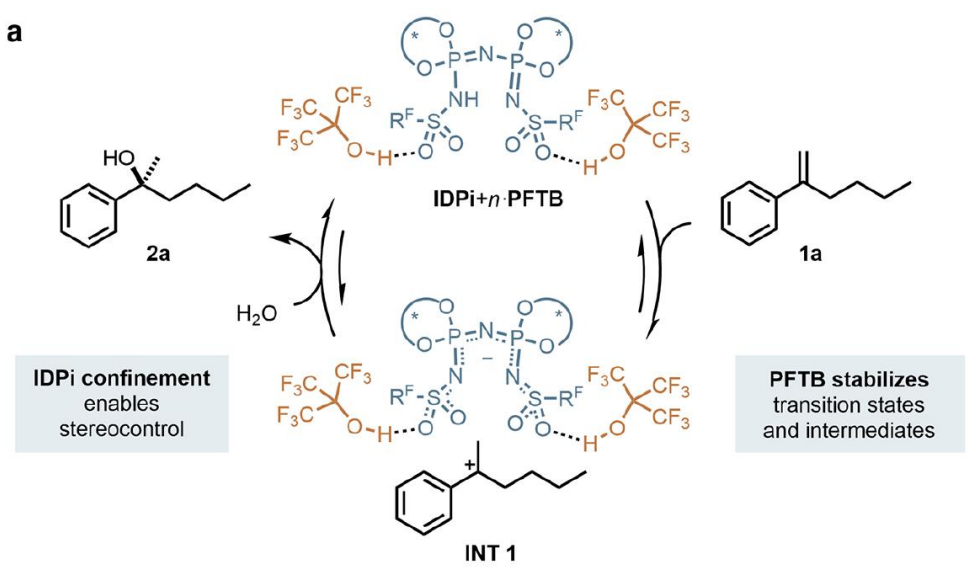


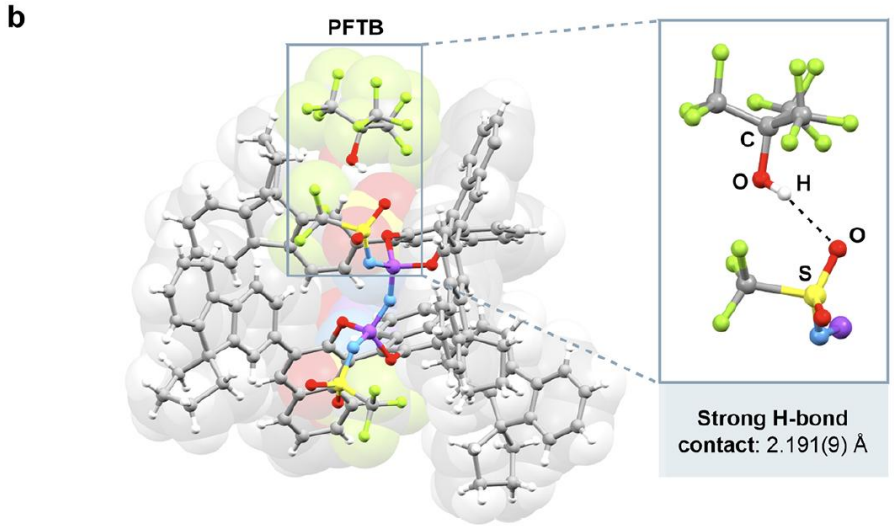
12 examples
up to 89% yield
up to 97:3 e.r.

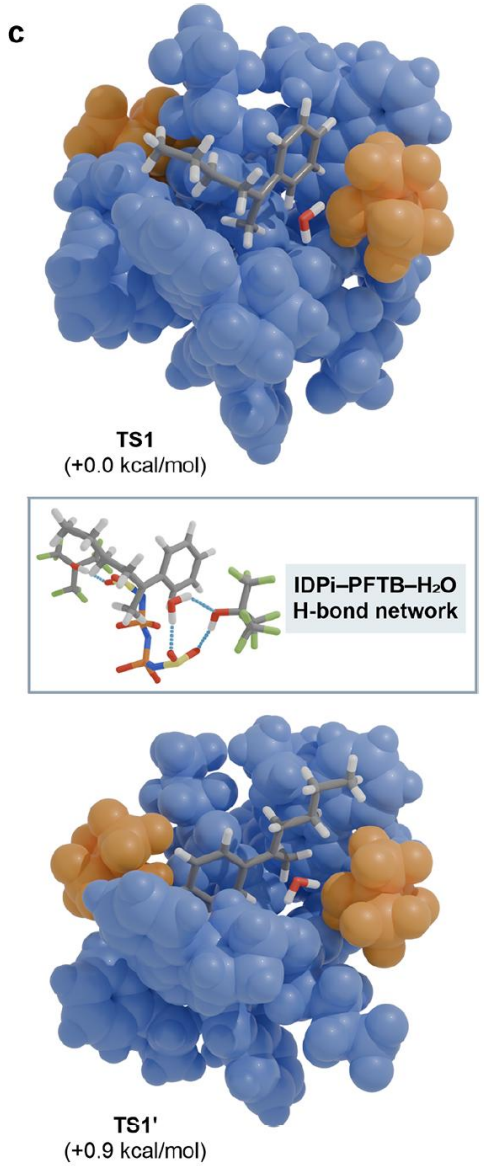


H. Hatano et al JACS, 2026, DOI: 10.1021/jacs.6c06916

Catalytic Asymmetric Hydration of Alkenes







H. Hatano et al JACS, 2026, DOI: 10.1021/jacs.6c06916

Teams

With Ukraine in my heart

Thank you for your attention!

ご清聴ありがとうございました！