RxnEnumProfiler

Virtual high-throughput screening of reaction networks

RxnEnumProfiler is a fully automated, massively parallel, out-of-the-box workflow developed to systematically evaluate large libraries of chemical species within a fixed reaction topology—that is, a predefined sequence of mechanistic steps involving reactants, products, intermediates, and/or transition states that characterize a catalytic or chemical process. Easily organize and manage all your data through a single, intuitive graphical interface—and obtain results in days or weeks instead of months or years.

Applications:

- Homogeneous (molecular) catalyst design
- Chemical (non-catalytic) reaction and reactivity optimization
- Minimizing unwanted reactions or improving selectivity

Key features:

- Multi-site R-group enumeration
- Fragment swapping
- Conformational sampling (Monte Carlo/ force fields or metadynamics/xTB)
- Exact and approximate (energetic span model) automated TOF calculations, % ee calculations, selectivity
- Quantum mechanical descriptor generation for cheminformatics ML
- Dataset creation for training machine learning potentials
- Optional thermochemical properties refinement, including anharmonic corrections and solvation entropy adjustments (e.g., Garza model)

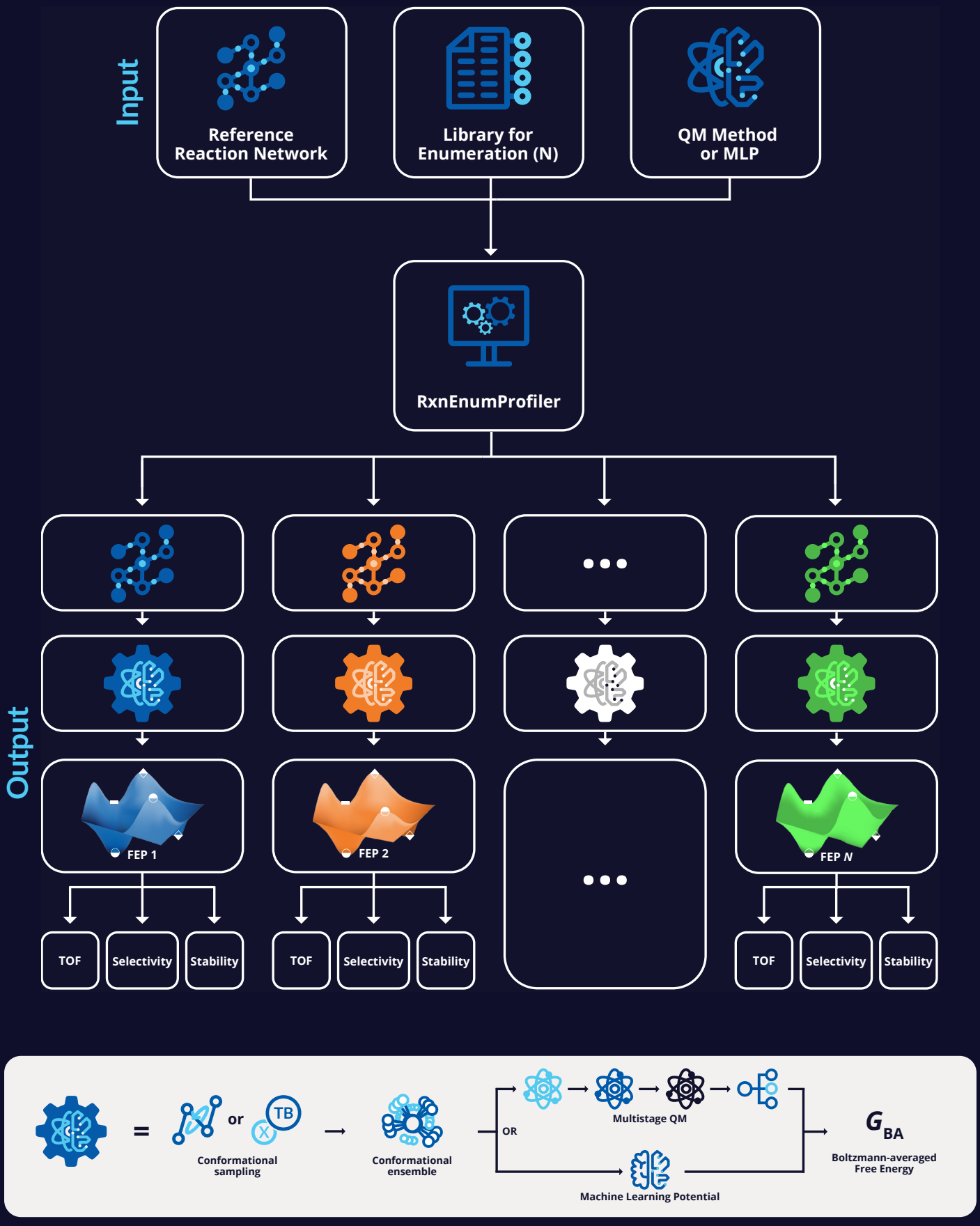
Methods:

- Classical force fields
- Extended tight-binding (GFN2-xTB)
- Pseudospectral density functional theory (PS-DFT)
- Machine learning potentials (e.g. MPNICE)

Schrödinger

RxnEnumProfiler

Virtual high-throughput screening of reaction networks



Schematic workflow of virtual high-throughput screening of reaction networks

Case example 1

Virtual high-throughput screening of enantioselectivity (ee) and turnover frequency (TOF) to guide the design of FLP-based asymmetric hydrogenation catalysts

Google Cloud

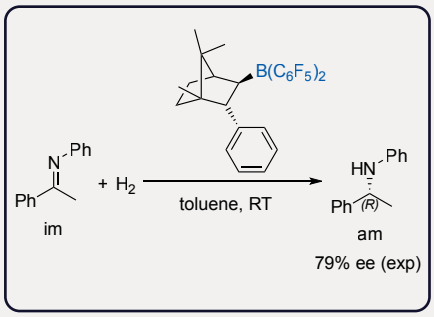
High Performance Computing (HPC) Cluster

Enatioselectivity screening:

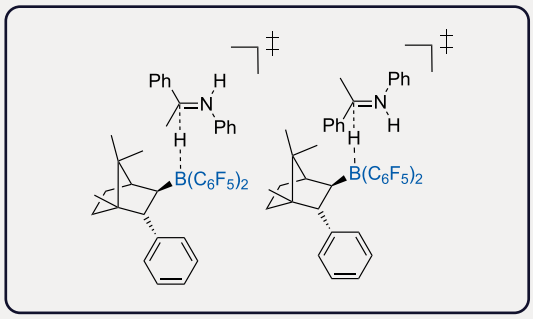
- 20 kcal/mol threshold for conformer selection
- ~400 transition states
- ~1600 DFT jobs
- 8 cores per DFT job
- ~2 weeks
- ~480,000 CPU hrs

TOF screening:

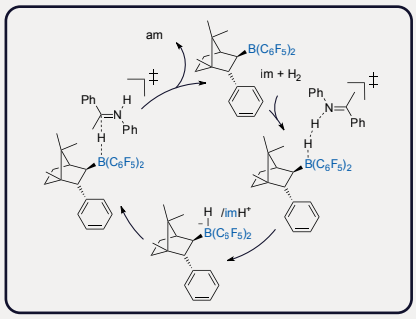
- Lowest conformer only
- 35 stationary points
- 10 transition states
- 175 DFT jobs
- 8 cores per DFT job
- ~5 days
- ~10,000 CPU hrs



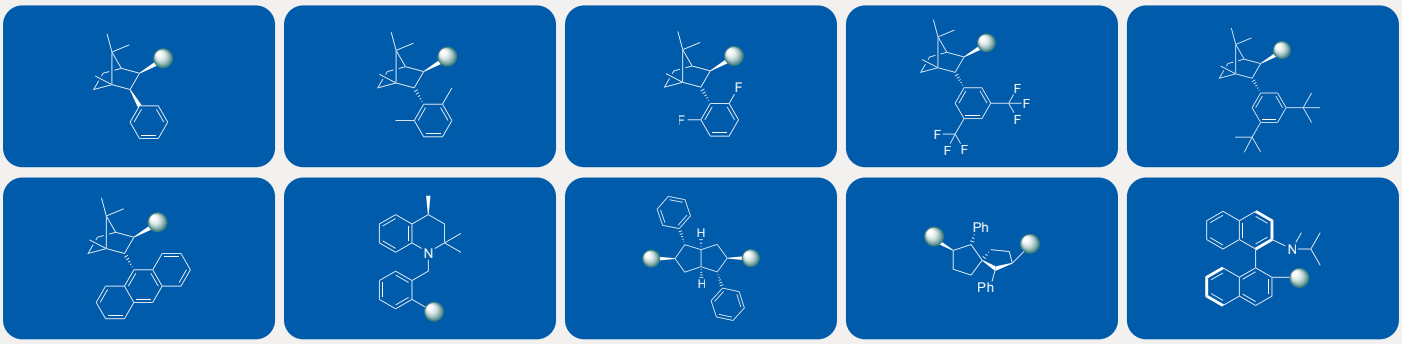
(a) asymmetric hydrogenation reaction



(b) reference network for enantioselectivity prediction



(c) reference network for TOF prediction



(d) library for single R-group enumeration (sphere indicates dummy atom)

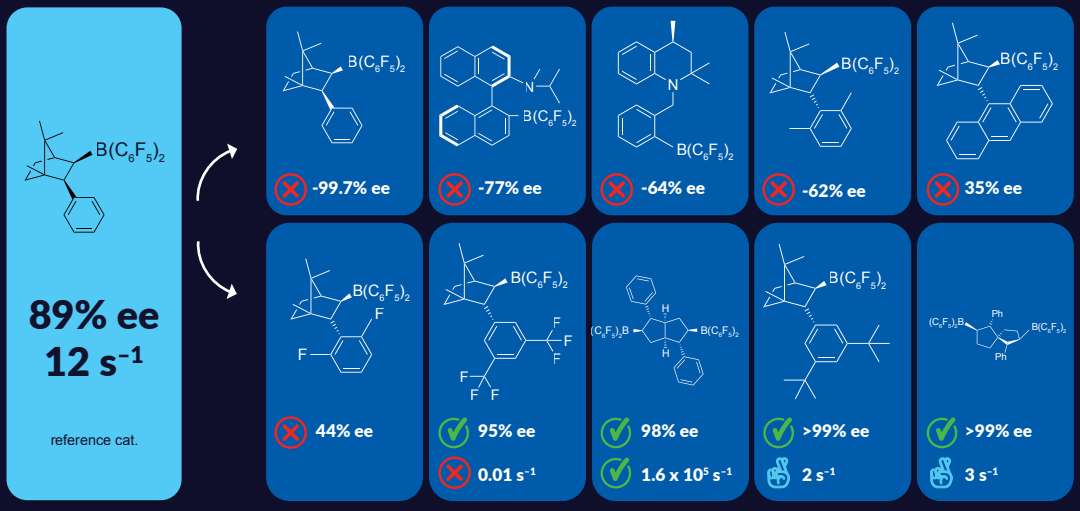
Predicted Boltzmann-averaged enantioselectivity (% ee) and turnover frequency (TOF in s^{-1})

ωB97X-D/C-PCM (toluene) with 6-311G++**//PS-6-31G*

UNSUITABLE

SUITABLE

POSSIBLE CANDIDATE



Case example 2

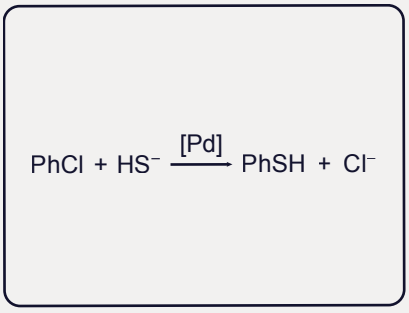
Virtual high-throughput screening of turnover frequency (TOF) to guide the design of Pd catalysts for C–O cross-coupling reaction

Google Cloud

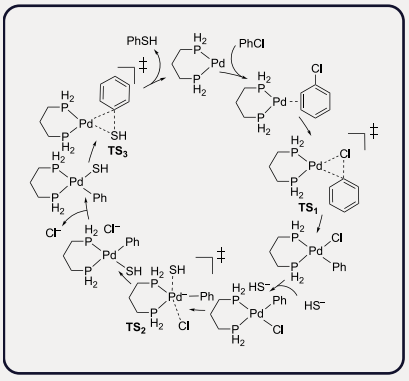
High Performance Computing (HPC) Cluster

- 10 conformers per stationary point
- 748 stationary points including
- 203 transition states
- 2992 DFT jobs

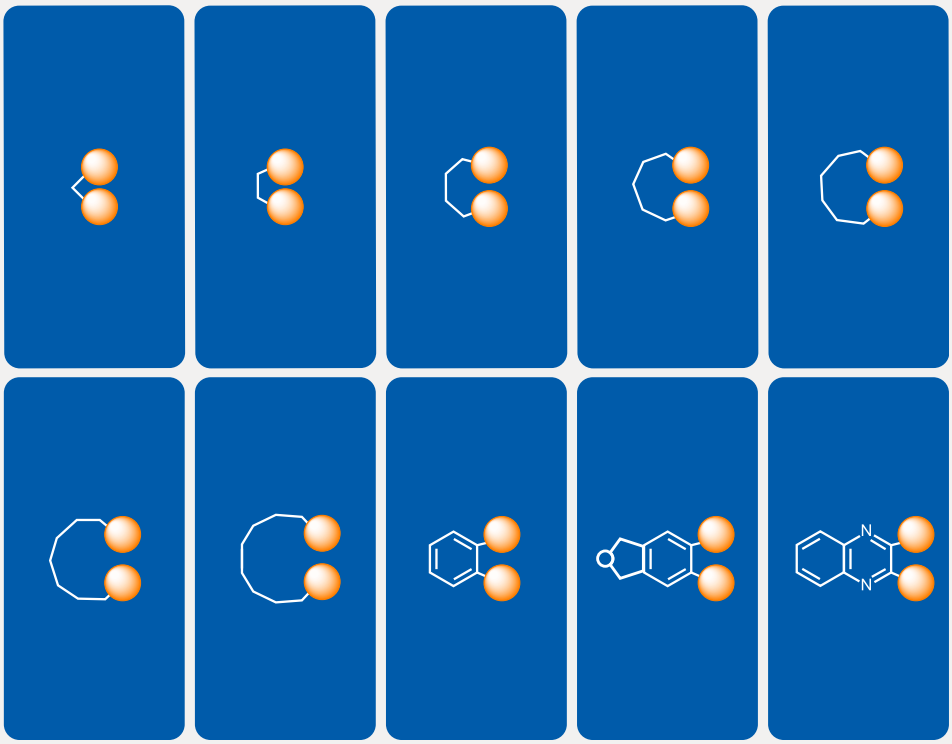
- 8 cores per DFT job
- ~2 days
- ~100,000 CPU hrs



(a) model C–O cross-coupling reaction



(b) reference network for TOF prediction

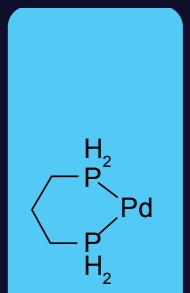


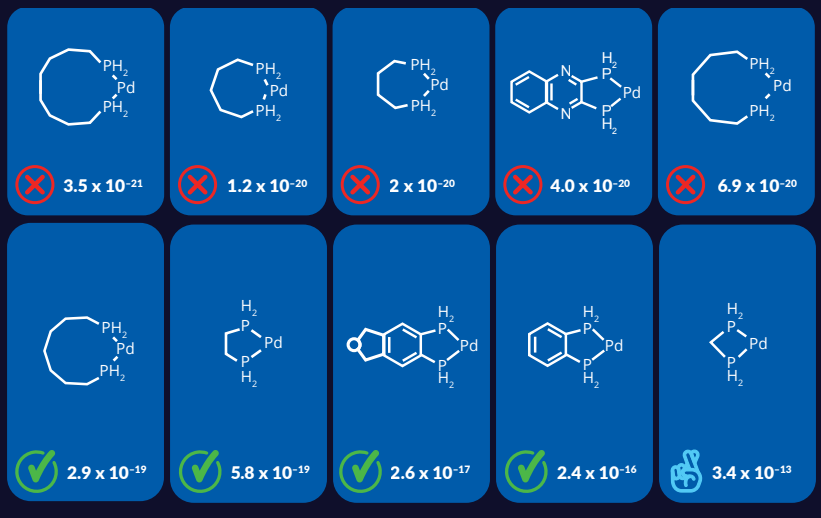
(c) virtual library for two R-group enumeration (spheres indicate dummy atoms)

Predicted Boltzmann-averaged turnover frequency (TOF in s^{-1})

B3LYP-D3 with PS-LACV3P++**//PS-LACVP*

- UNSUITABLE
- SUITABLE
- POSSIBLE CANDIDATE (POSSIBLE CHANGE IN MECHANISM DUE TO PARTIAL PP LIGAND DISSOCIATION)





Case example 3

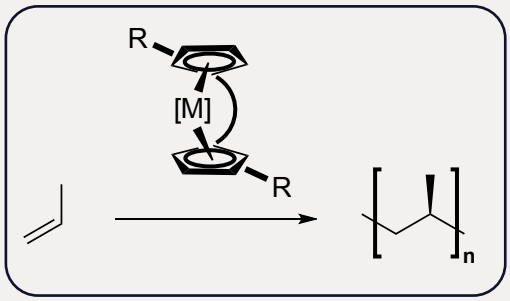
Designing an ansa-metallocene catalyst for enantiopure isotactic propylene polymerization, with the goals of minimizing stereochemical and regioselectivity errors while maximizing molecular weight and reaction rate

Google Cloud

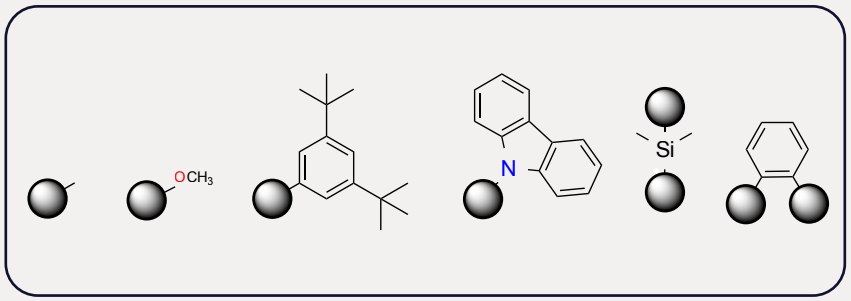
High Performance Computing (HPC) Cluster

- Lowest conformer only
- 55 stationary points
- 44 transition states
- 220 DFT jobs

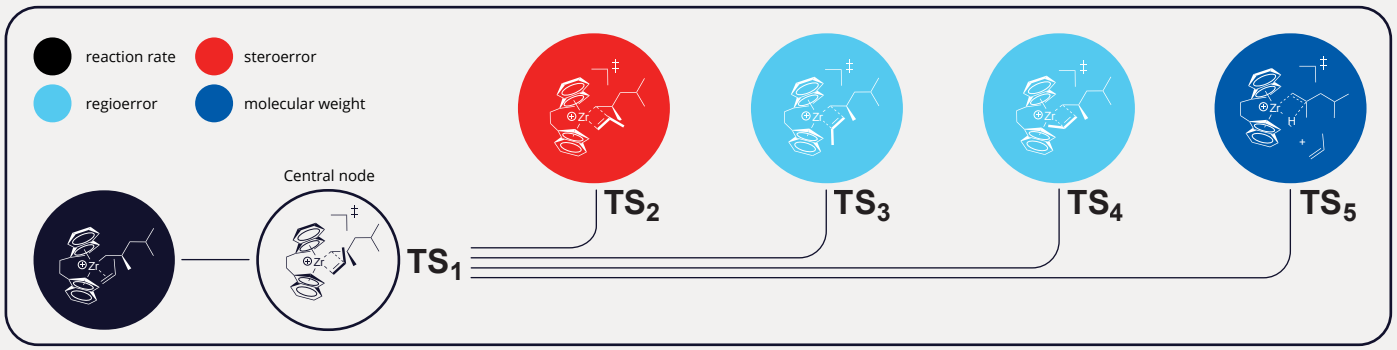
- 8 cores per DFT job
- ~2 days
- ~8,800 CPU hrs



(a) enantiopure isotactic propylene polymerization



(b) library for enumeration (sphere indicates dummy atom)

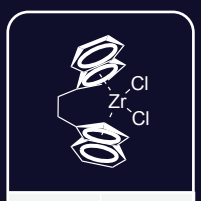
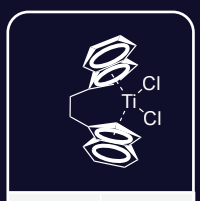
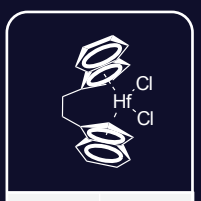
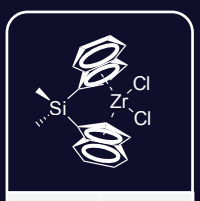
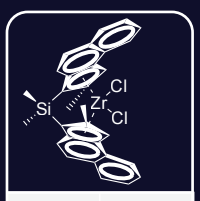
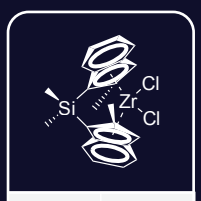
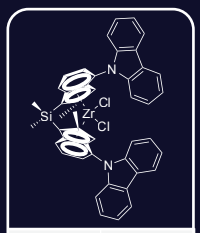
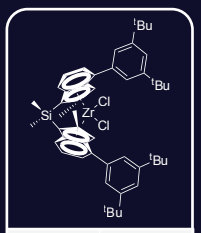
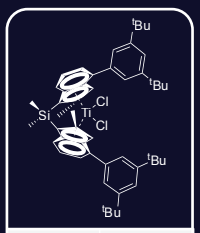
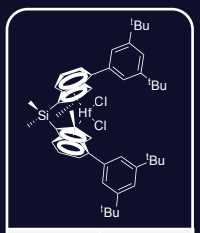
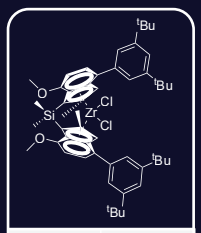


(c) reference network for catalyst design

Score functions for:

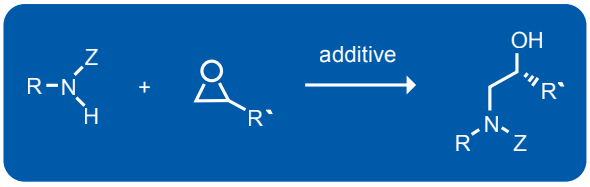
B3LYP-D3 with
PS-LACV3P++**//PS-LACVP*

- steroerror
- regioerror
- molecular weight
- reaction rate

					
1.1	2.8	3.1	3.3	4.0	4.5
1.4	3.6	2.6	3.2	3.4	3.9
−7.3	−1.9	−3.4	−2.4	−0.2	−2.3
1.3×10^9	2.7×10^6	3.4×10^5	3.6×10^5	1.5×10^7	1.3×10^7
					
10	10	10.2	10.4	11.3	
6.1	7.2	8.2	6.1	7.2	
5.2	3.8	6.0	6.3	2.2	
1.8×10^5	1.3×10^7	1.7×10^5	2.8×10^4	5.1×10^6	

Case example 4

Virtual high-throughput screening of rate constant for epoxy-amine curing reaction



Additive: no additive, BF_3, MeOH, H_2O

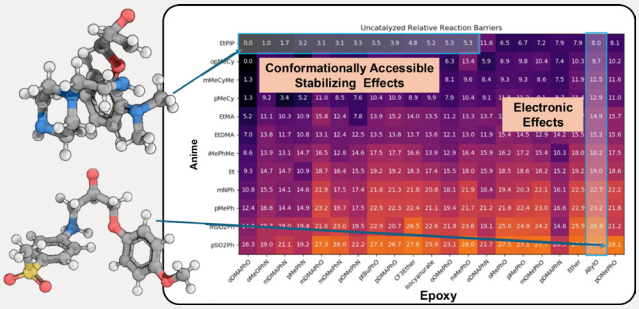
R = 12 Fragments

R' = 21 Fragments

Z = H or R

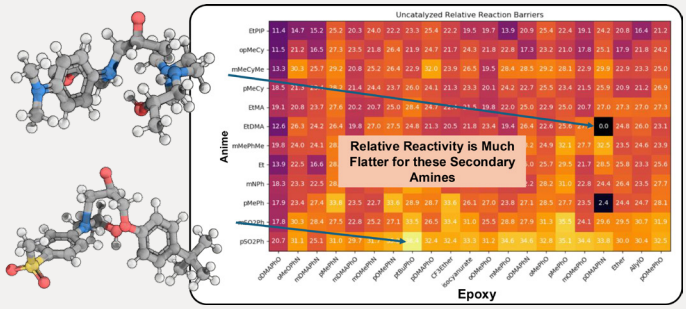
- 252 Epoxy/Amine combinations
- ~240 conformations per Epoxy/Amine combination
- 2 DFT calculations per conformation
- Primary and Secondary reaction pathways
- 2 RxnEnumProfiler jobs
- 504 RxnProfiler subjobs
- ~480,000 DFT calculations

Primary Amine Uncatalyzed Relative Reaction Barriers



All geometries and energies B3LYP-D3/LACVP** (kcal/mol)

Secondary Anime Uncatalyzed Relative Reaction Barriers



All geometries and energies B3LYP-D3/LACVP** (kcal/mol)

- Relative reactivity for secondary amines flatter than primary amines due to included hydroxyl group for stabilization
- Derivatives with hydrogen donating/accepting moieties can drop activation energy
- Amines reactivity strongly influenced by electronics
  - EWG slower
  - EDG faster
- Reactivity – BF_3 > MeOH > H_2O >> Uncatalyzed
- High throughput reaction screening provides the data necessary for machine learning

Contact us: ms-sales@schrodinger.com

Schrödinger

Learn more about RxnEnumProfiler