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 force fields
- Optional thermochemical properties refinement, including anharmonic corrections and solvation entropy adjustments (e.g., Garza model)

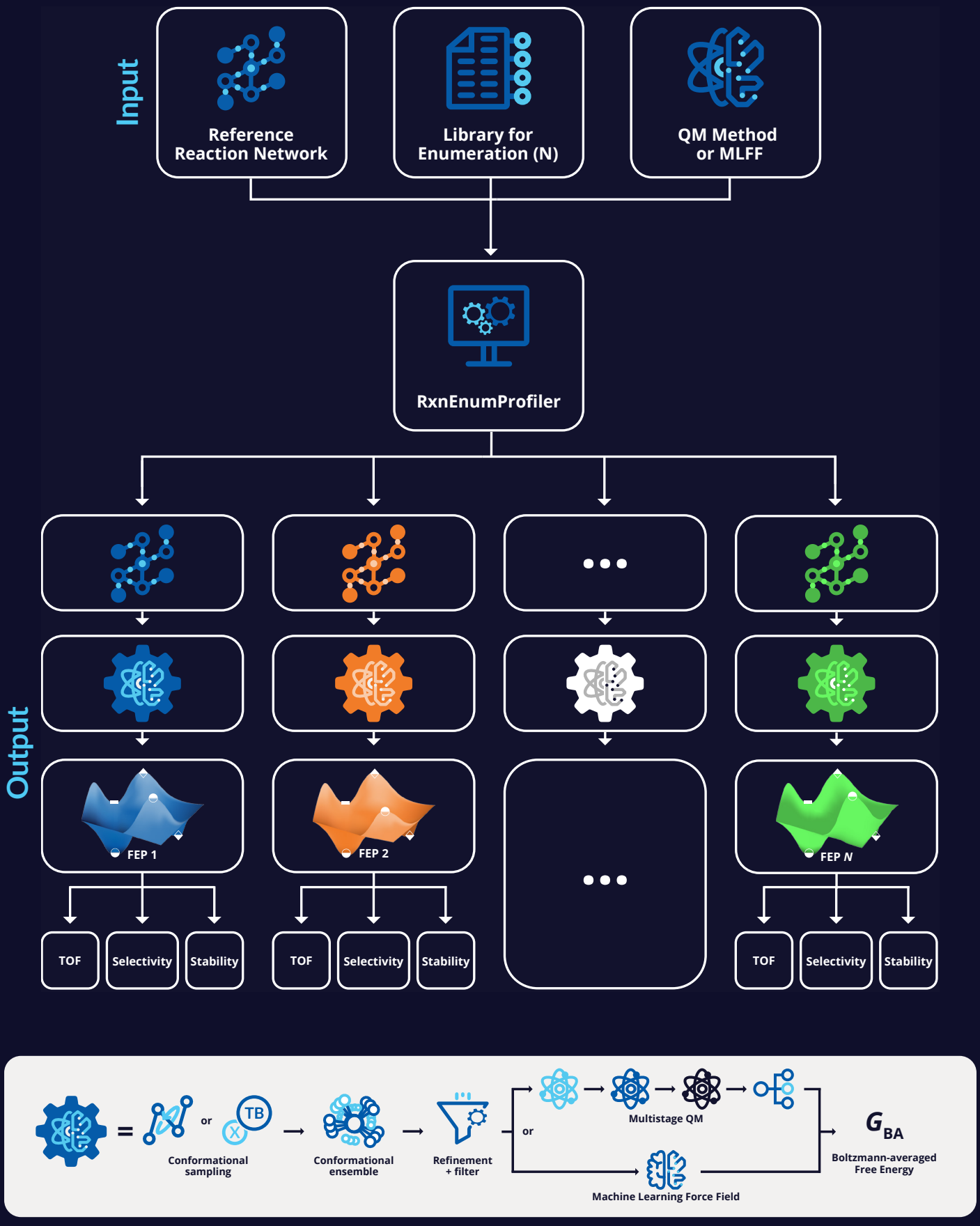
Methods:

- Classical force fields
- Extended tight-binding (GFN2-xTB)
- Pseudospectral density functional theory (PS-DFT)
- Machine Learning Force Fields (e.g. MPNICE, UMA_SM_OMOL)

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

Enantioselectivity screening:

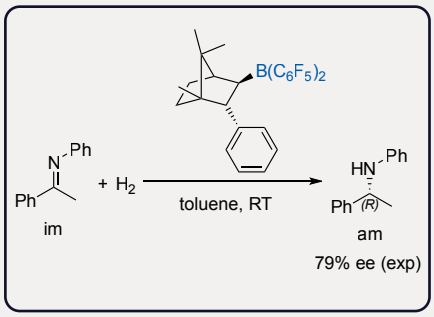
- 20 kcal/mol threshold for conformer selection
- ~400 transition states
- ~1600 DFT jobs

- 8 cores per DFT job
- ~2 weeks
- ~105,000 CPU hrs

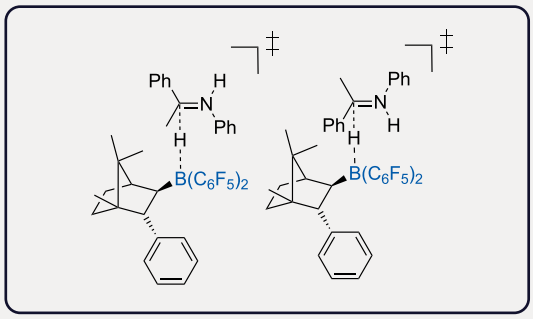
TOF screening:

- Lowest conformer only
- 35 stationary points
- 10 transition states
- 175 DFT jobs

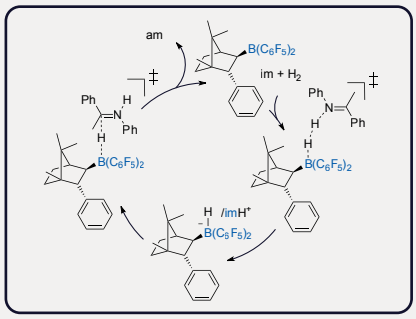
- 8 cores per DFT job
- ~5 days
- ~4,800 CPU hrs



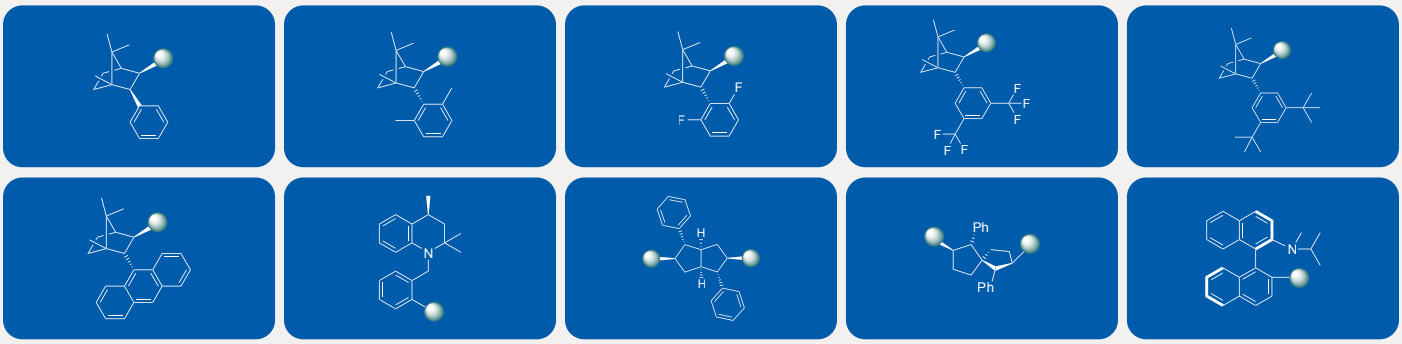
(a) asymmetric hydrogenation reaction



(b) reference network for enantioselectivity prediction



(c) reference network for TOF prediction



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

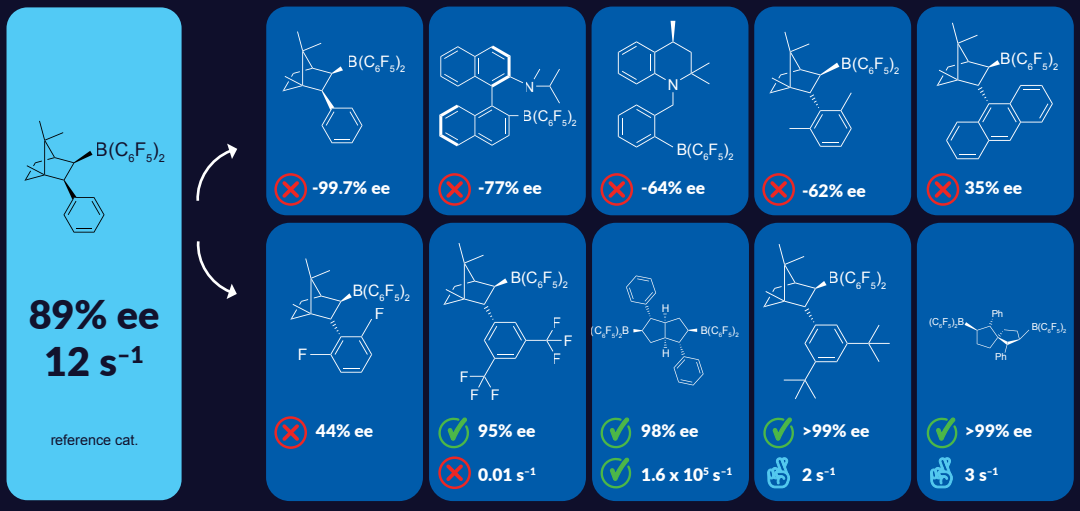
Predicted Boltzmann-averaged enantioselectivity (% ee) and turnover frequency (TOF in s^{-1})

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

UNSUITABLE

SUITABLE

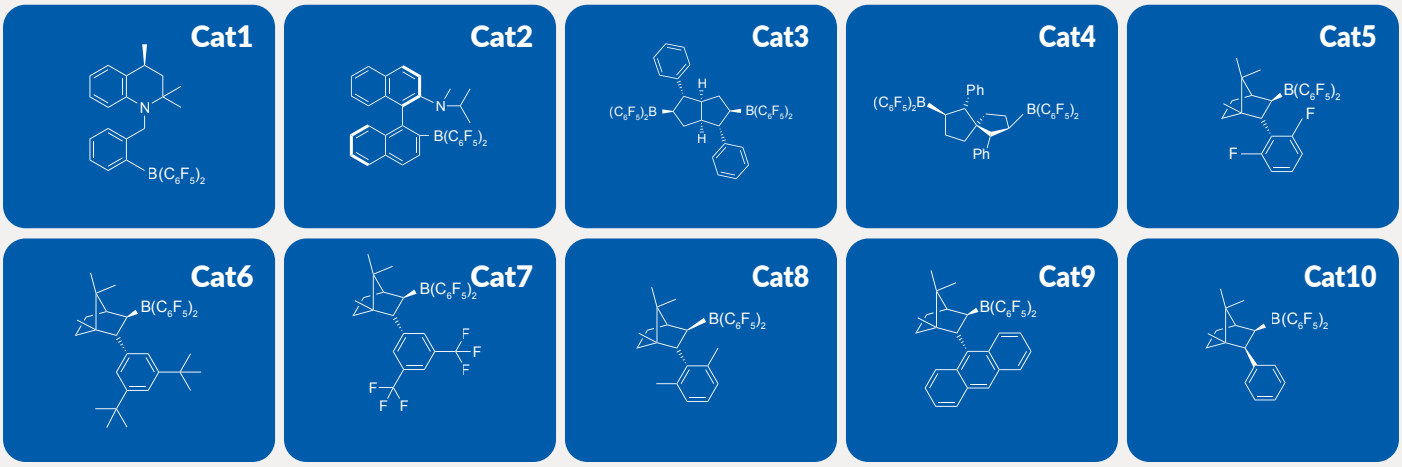
POSSIBLE CANDIDATE



Case example 1

DFT vs Machine Learning Force Fields for Catalyst Design: Accuracy–Efficiency Tradeoffs

394 cores requested on queuing system for MLFF
3152 cores used in parallel for DFT



Physics

Machine learning

Method	Established Trend	Failures	Time
ωB97X-D/C-PCM (toluene) with 6-311G++**// PS-6-31G*	cat10 < cat2 < cat1 < cat8 < cat9 < cat5 < ref < cat7 < cat3 < cat6 < cat4 -99.7% -77% -64% -62% 35% 44% 89% 95% 98% >99% >99%	33 of the 394 jobs failed	~ 2 weeks, ~105,000 CPU hrs
UMA_SM_ OMOL v1.1	cat10 < cat2 < cat8 < cat1 < cat5 < cat9 < ref < cat7 < cat6 < cat4 < cat3 -99.6% -95% -72% -62% 35% 44% 89% 95% 98% >99% >99%	4 of the 394 jobs failed	~ 8 hrs, ~1,100 CPU hrs
Organic_ MPNICE_tb (toluene)	cat2 < cat10 < cat1 < cat8 < cat7 < cat5 < cat9 < cat3 < ref < cat4 < cat6 -99.5% -98% -77% 10% 23% 27% 78% 92% 93% 99.5% 99.6%	24 of the 394 jobs failed	~ 9 hrs, ~360 CPU hrs

Case example 2

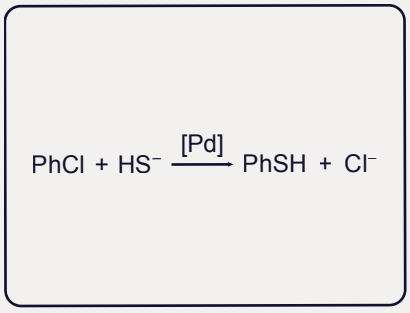
Virtual high-throughput screening of turnover frequency (TOF) to guide the design of Pd catalysts for C–O cross-coupling reaction

Google Cloud

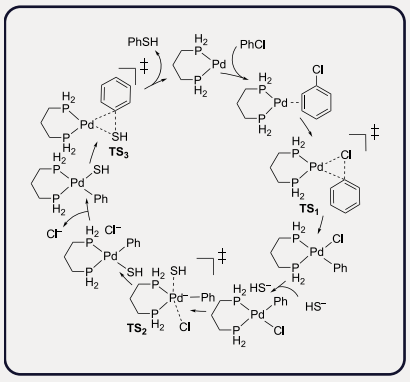
High Performance Computing (HPC) Cluster

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

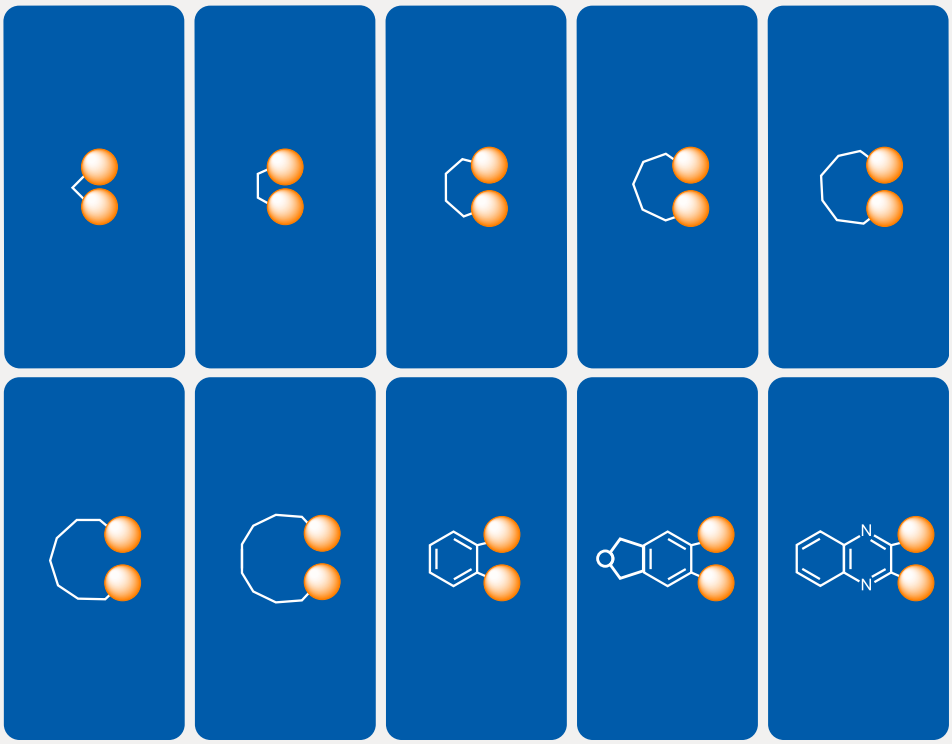
- 8 cores per DFT job
- ~2 days
- ~3,500 CPU hrs



(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)

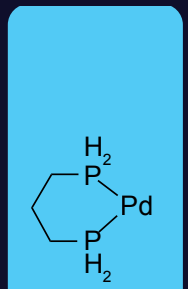
Predicted Boltzmann-averaged turnover frequency (TOF in s^{-1})

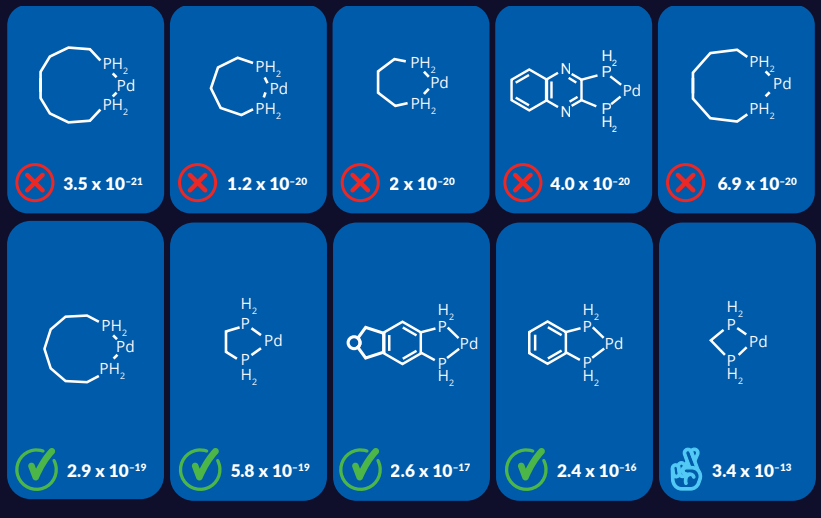
B3LYP-D3 with PS-LACV3P++**//PS-LACVP*

UNSUITABLE

SUITABLE

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





Case example 3

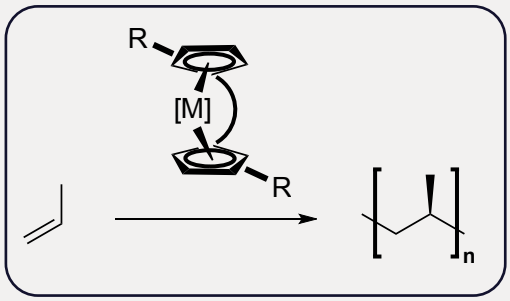
Designing an ansa-metallocene catalyst for enantiopure isotactic propylene polymerization, with the goals of minimizing stereochemical and regioselectivity errors while maximizing molecular weight and reaction rate

Google Cloud

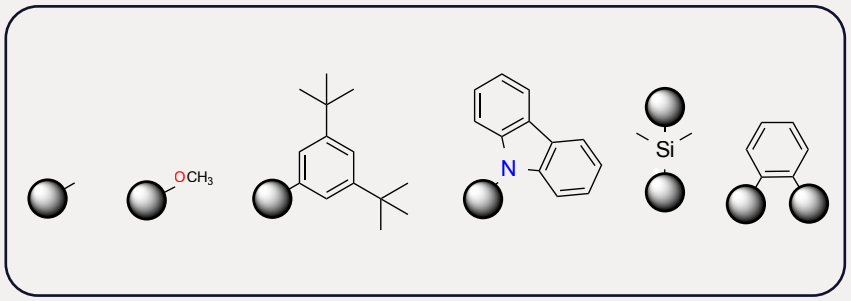
High Performance Computing (HPC) Cluster

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

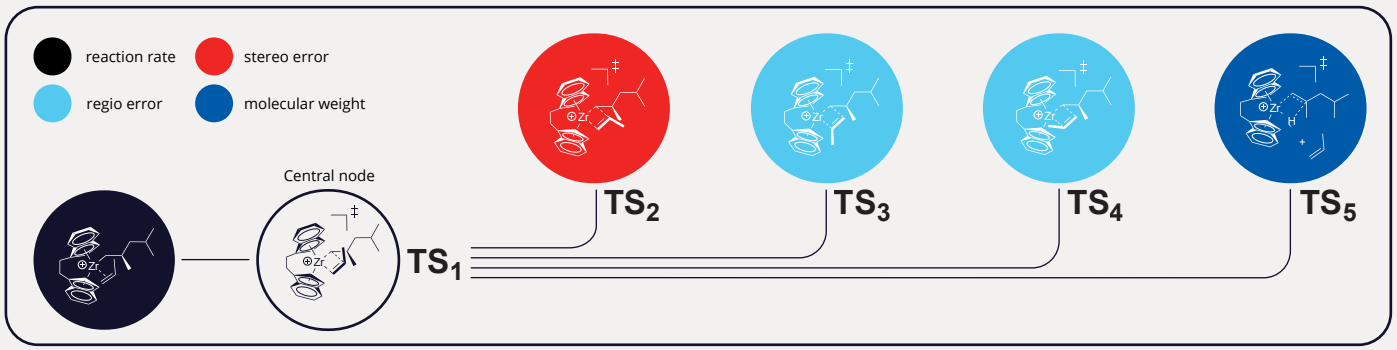
- 8 cores per DFT job
- ~2 days
- ~4,100 CPU hrs



(a) enantiopure isotactic propylene polymerization



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

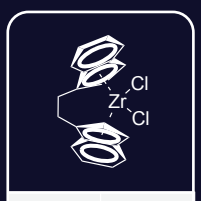
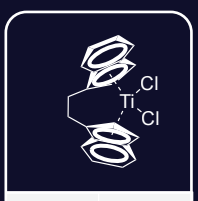
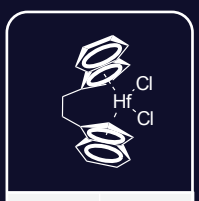
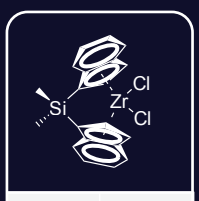
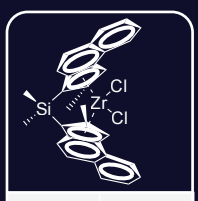
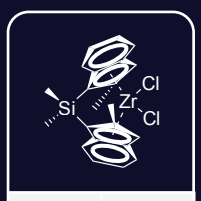
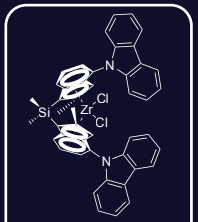
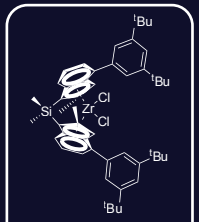
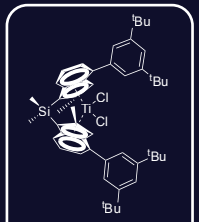
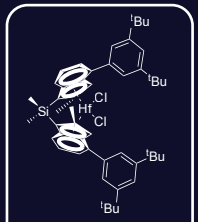
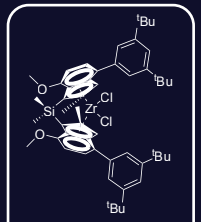


(c) reference network for catalyst design

Score functions for:

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

- stereo error
- regio error
- molecular weight
- reaction rate

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

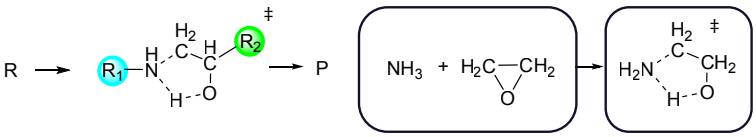
Case Example 4

Screen substituents on the amine and epoxide in the epoxy–amine curing reaction and rank reactivity based on activation barrier

Computational Setup & Scale

- 2500 reactions, 5000 stationary points (2500 transition states)
- Macromodel engine, OPLS4 force field, lowest conformer
- 5000 CPU cores requested (queuing system)
- UMA_SM_OMOL MLFF v 1.1
- 2489 of 2500 job(s) succeeded
- ~3 hrs walltime, ~500 CPU hrs

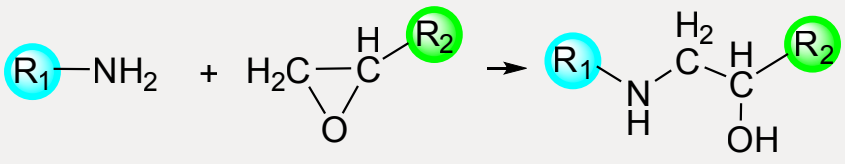
Concerted Channel and Reference Reaction Network



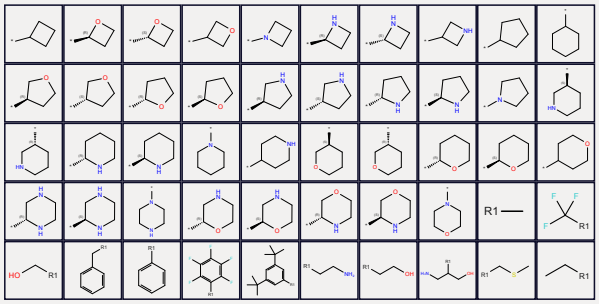
Reactivity Heatmap & Barriers

Concerted Channel: barriers span from 25 to 67 kcal/mol

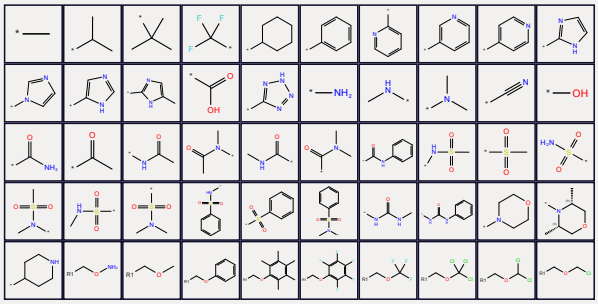
Reaction Chemistry & Libraries



R_1-library



R_2-library



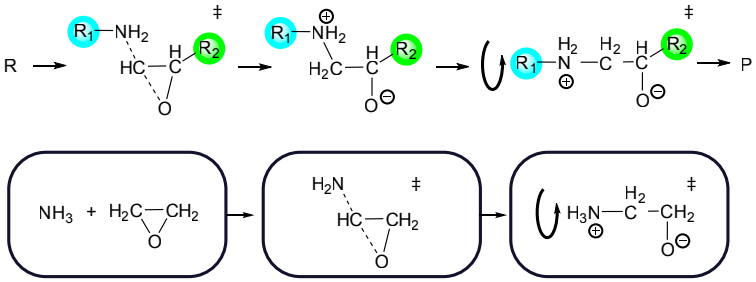
Case Example 4

Screen substituents on the amine and epoxide in the epoxy–amine curing reaction and rank reactivity based on activation barrier

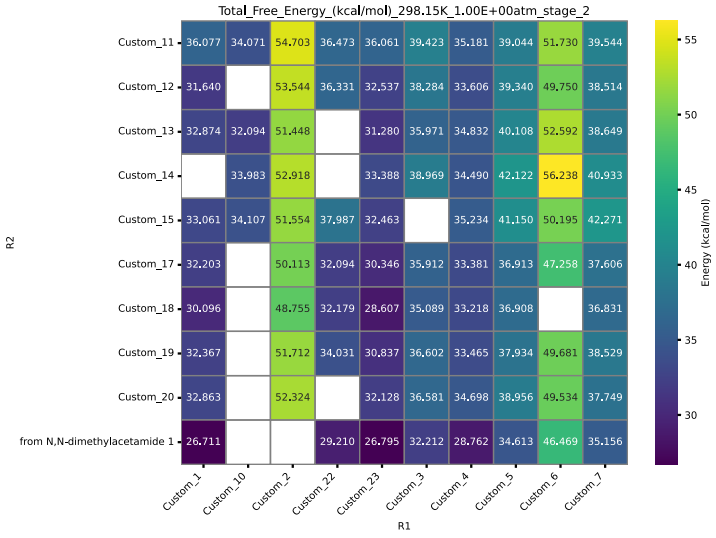
Computational Setup & Scale

- 2500 reactions, 7500 stationary points (5000 transition states)
- Macromodel engine, OPLS4 force field, lowest conformer
- 7500 CPU cores requested (queuing system)
- UMA_SM_OMOL MLFF v 1.1
- 2142 of 2500 job(s) succeeded
- ~3.5 hrs walltime, ~1000 CPU hrs

Stepwise Channel and Reference Reaction Network



Reactivity Heatmap & Barriers



Stepwise Channel: barriers span from 20 to 58 kcal/mol

Contact us: ms-sales@schrodinger.com

Schrödinger

Learn more about RxnEnumProfiler