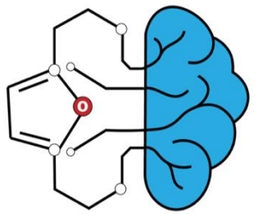
DC3 – Karoline Schjelde

Predicting side reactions using a combined meta-dynamics-MD and ML approach

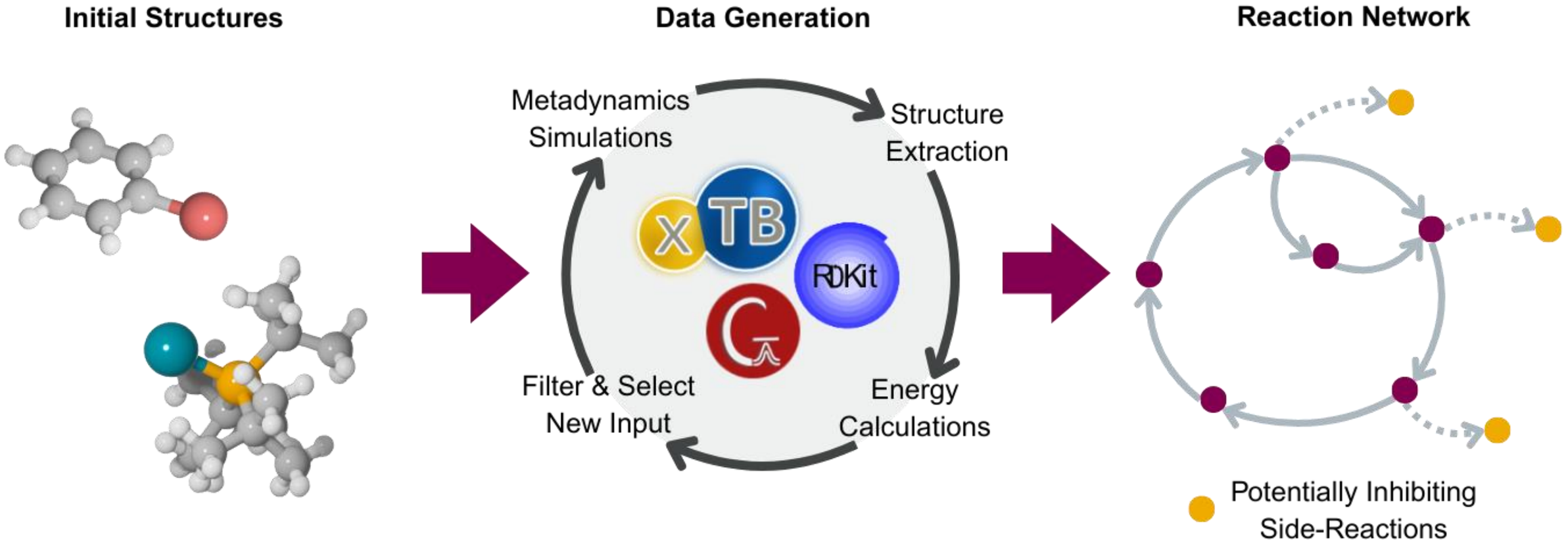
Academic Supervisor: Jan Halborg Jensen | Industry Supervisor: Mikhail Kabeshov



UNIVERSITY OF COPENHAGEN

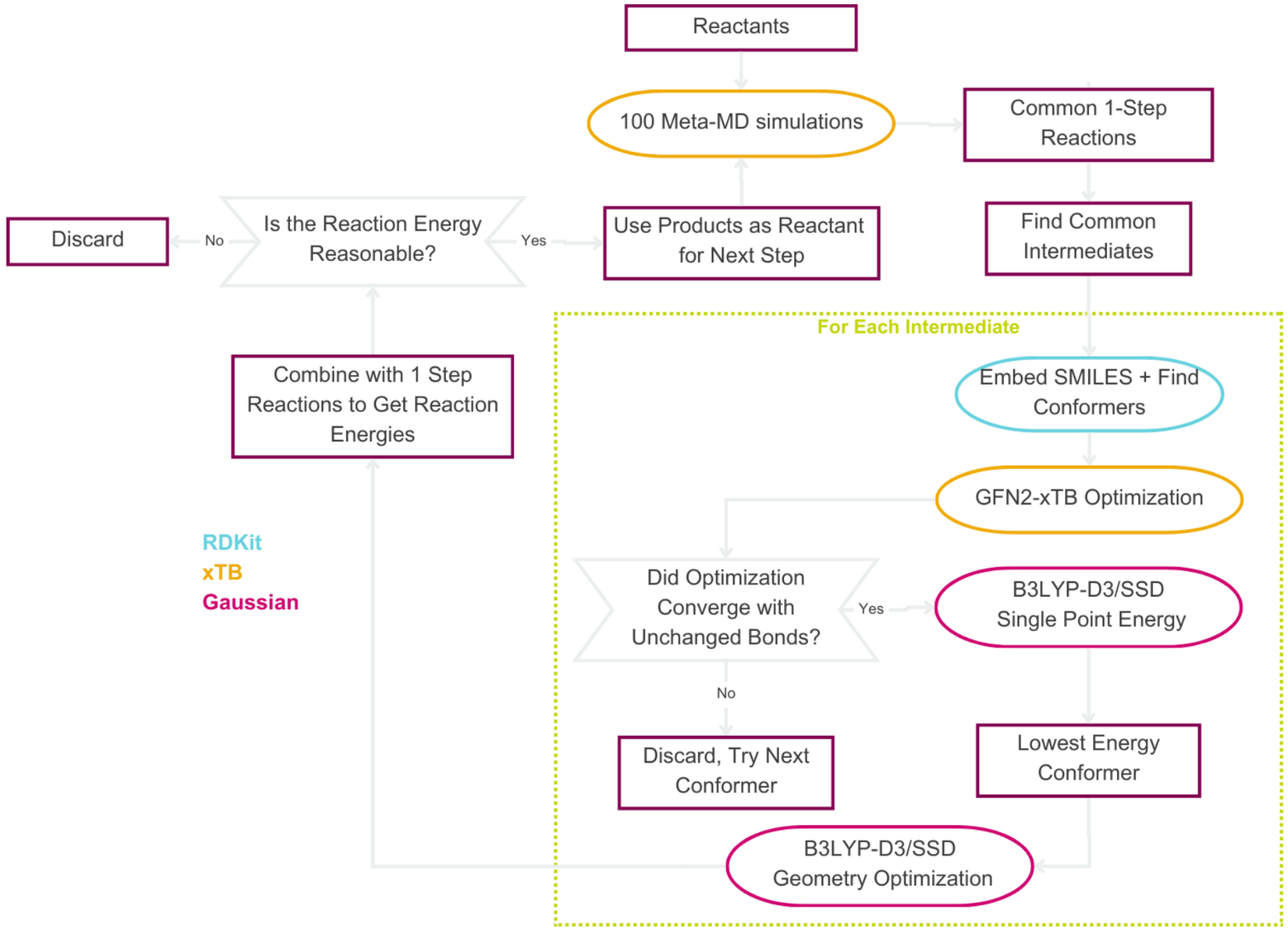
AstraZeneca

Overview


Initial Structures
Data Generation
Reaction Network
Metadynamics Simulations
Structure Extraction
Filter & Select New Input
Energy Calculations
xTB
RDKit
Potentially Inhibiting Side-Reactions

Overview

- What I am currently working on
- Implemented using python and SLURM
- Should be automatic eventually..



UCPH

Overview

- Automatic generation of the catalytic cycle as PoC

Role of the Base in Buchwald–Hartwig Amination

Ylva Sunesson,[†] Elaine Limé,[‡] Sten O. Nilsson Lill,[‡] Rebecca E. Meadows,[§] and Per-Ola Norrby*[,†,‡]

[†]Department of Chemistry and Molecular Biology, University of Gothenburg, Kemigården 4, SE-412 96 Göteborg, Sweden
[‡]Pharmaceutical Development, Global Medicines Development, AstraZeneca, Pepparedsleden 1, SE-431 83 Mölndal, Sweden
[§]Pharmaceutical Development, AstraZeneca, Silk Road Business Park, Macclesfield SK10 2NA, United Kingdom

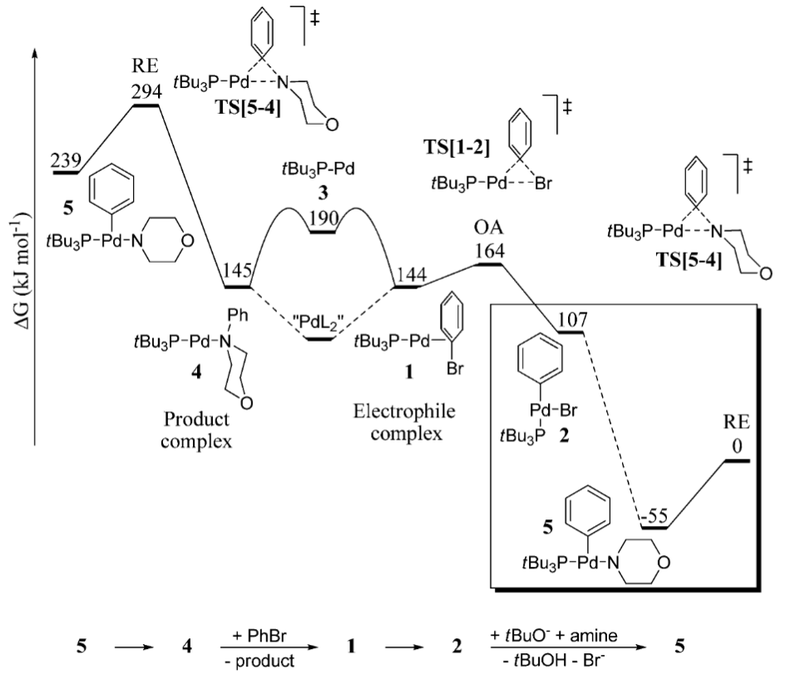


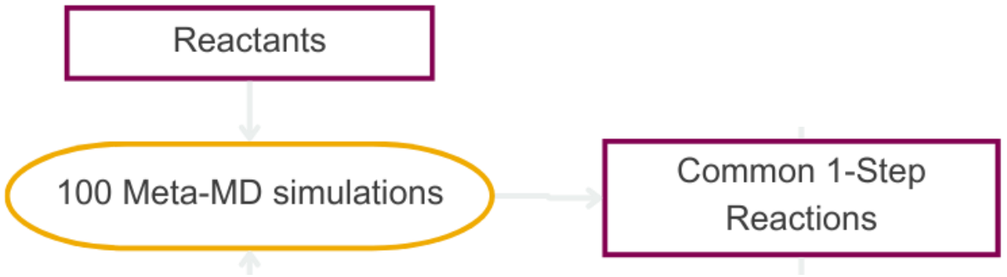
Figure 3. Free energy profile (FEP) of the catalytic cycle in benzene. Curved lines represent monotonous reactions without a potential energy barrier, that is, a diffusion-controlled free energy barrier of ~20 kJ mol^{-1} to association.[22] Dotted lines show multistep reactions with expected low barriers, but potentially including low-energy intermediates. The boxed area includes multiple steps that we investigated further (*vide infra*).

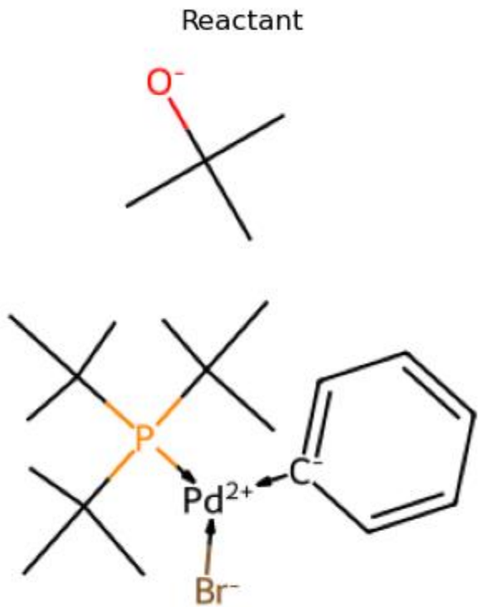
Meta-MD simulations

- Run with two sets of params:
  - *Runs until bond breaks*

    Looking at simulation time steps, gathering reaction steps for each

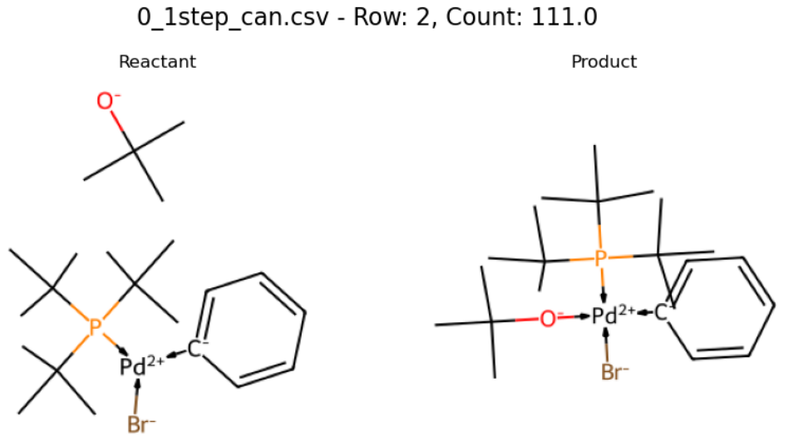
- Takes SMILES as input





1 Step Reactions

- Sorted based on canonical smiles:
  - *Counts occurrence*
    - Two types
  - *Saves SMILES*
  - *Canonical smiles and metal centers -> issue*
    - Random



	Reactants_canonical	Products_canonical	Reactions	reactant_smiles_am	product_smiles_am	counts
0	CC(C)(C)P(->[Pd+2](<-[Br-])<-[c-]1ccccc1)(C(C)(C)C)C(C)(C)C.CC(C)(C)[O-]	CC(C)(C)P1(C(C)(C)C)->[Pd				
1	CC(C)(C)P(->[Pd+2](<-[Br-])<-[c-]1ccccc1)(C(C)(C)C)C(C)(C)C.CC(C)(C)[O-]	CC(C)(C)[O-]->[Pd+2](<-[B				
2	CC(C)(C)P(->[Pd+2](<-[Br-])<-[c-]1ccccc1)(C(C)(C)C)C(C)(C)C.CC(C)(C)[O-]	CC(C)(C)[O-]->[Pd+2]1(<-[				
3	CC(C)(C)P(->[Pd+2](<-[Br-])<-[c-]1ccccc1)(C(C)(C)C)C(C)(C)C.CC(C)(C)[O-]	CC(C)([O-])C[H]->[Pd+2](<				
4	CC(C)(C)P(->[Pd+2](<-[Br-])<-[c-]1ccccc1)(C(C)(C)C)C(C)(C)C.CC(C)(C)[O-]	CC(C)([O-])C[H]->[Pd+2]1(				

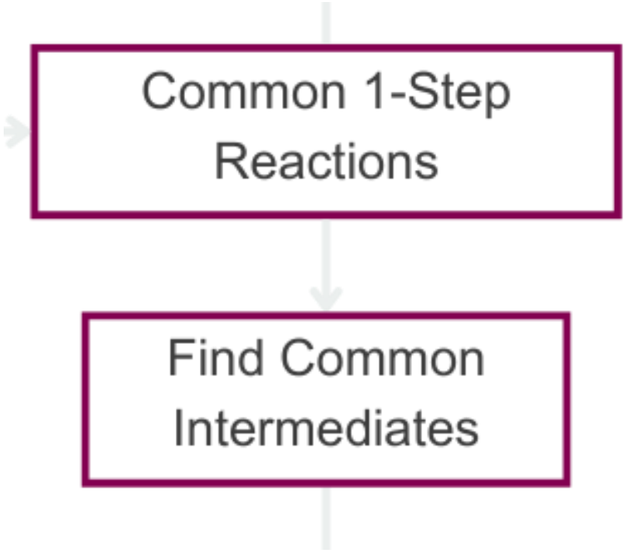
UCPH

Common Intermediates

- Unique Molecules:
  - *One mol at a time*

    Avoid doing the same calculations over and over

	Unique_SMILES
0	C=C(C)C
1	CC(C)(C)P(->[Pd+2](<-[Br-])<-[c-]1ccccc1)(C(C)(C)C)C(C)(C)C
2	CC(C)(C)[O-]->[Pd](<-[Br-])(<-C1=CC=C=CC=1)<-P(C(C)(C)C)(C(C)(C)C)C(C)(C)C
3	CC(C)(C)P1(C(C)(C)C)->[Pd+2](<-[Br-])(<-[c-]2ccccc2)<-[H]CC1(C)C
4	CC(C)(C)[O-]->[Pd+2]1(<-[Br-])(<-[c-]2ccccc2)<-[H]CC(C)(C)P->1(C(C)(C)C)C(C)(C)C
5	CC(C)(C)Oc12->[Pd](<-[Br-])(<-P(C(C)(C)C)(C(C)(C)C)C(C)(C)C)<-c1cccc2
6	CC(C)(C)OP(C(C)(C)C)(C(C)(C)C)->[Pd+2](<-[Br-])(<-[c-]1ccccc1)<-[C-](C)(C)C
7	CC(C)(C)[O-]->[Pd+2](<-[Br-])(<-[c-]1ccccc1)<-P(C(C)(C)C)(C(C)(C)C)C(C)(C)C
8	CC1(C)C[H]->[Pd](<-[Br-])(<-[c-]2ccccc2)<-P(C(C)(C)C)(C(C)(C)C)O1
9	CC(C)(C)[O-]
10	CC(C)([O-])C[H]->[Pd+2](<-[Br-])(<-[c-]1ccccc1)<-P(C(C)(C)C)(C(C)(C)C)C(C)(C)C
11	CC(C)(C)[O-]->[Pd+2]12(<-[Br-])(<-P(C(C)(C)C)(C(C)(C)C)C(C)(C)C)<-C3=C->1[CH-]->2C=C=C3
12	CC(C)(C)[O-]->[Pd+2]1(<-[Br-])(<-[c-]2ccccc2)<-[H]CC(C)(C)P->1C(C)(C)C
13	[H+]

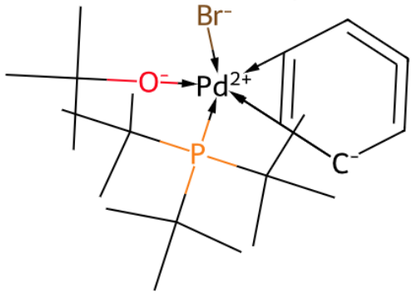


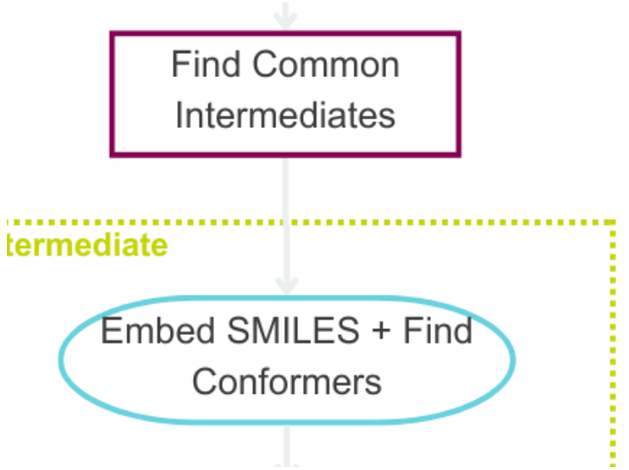
UCPH

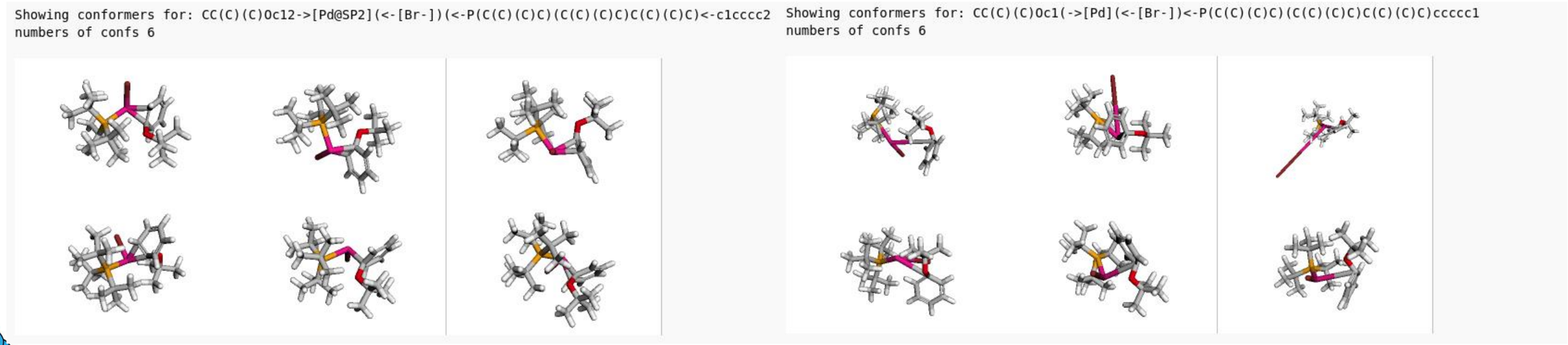
SMILES to XYZ

- RDKit:
  - *Embedding comes with its own set of issues*
    - Changing two dative to one
    - Planer configurations around metal
    - Adding Hydrogens
  - *Conformers*
    - 3 + 3 * Rotatable bonds

Problem! >:(

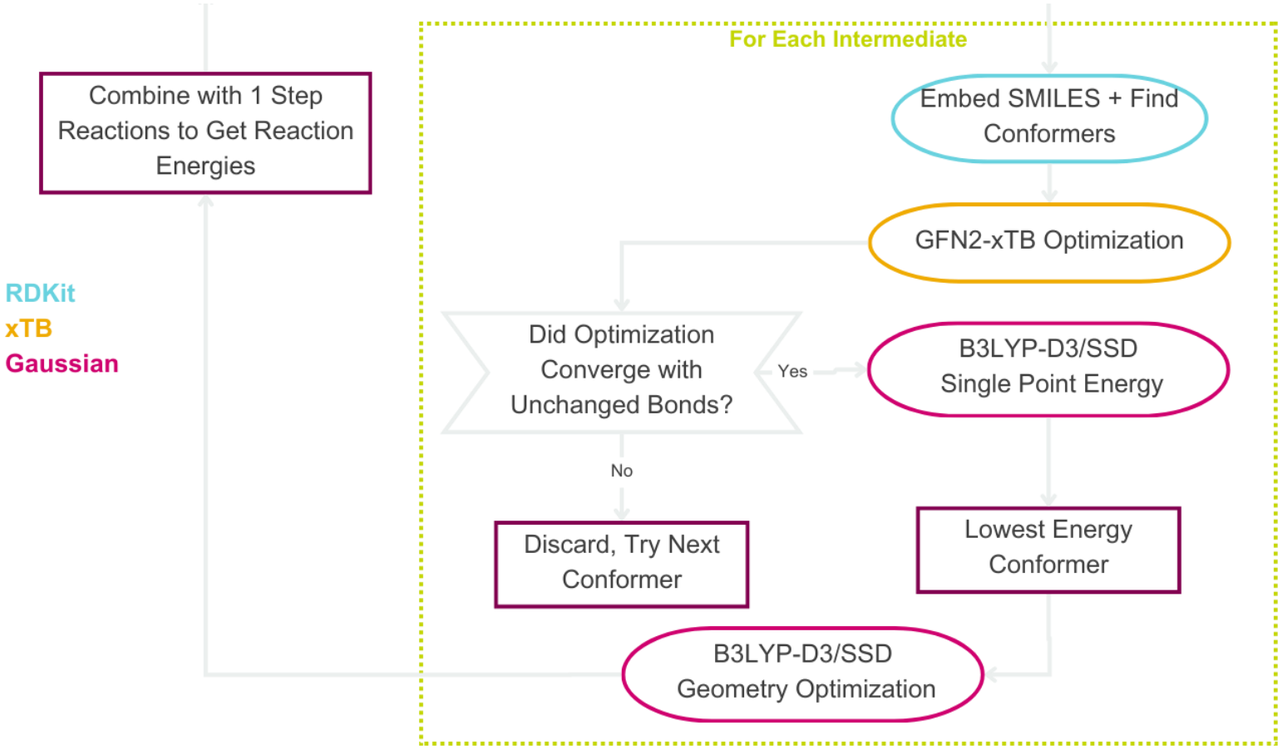






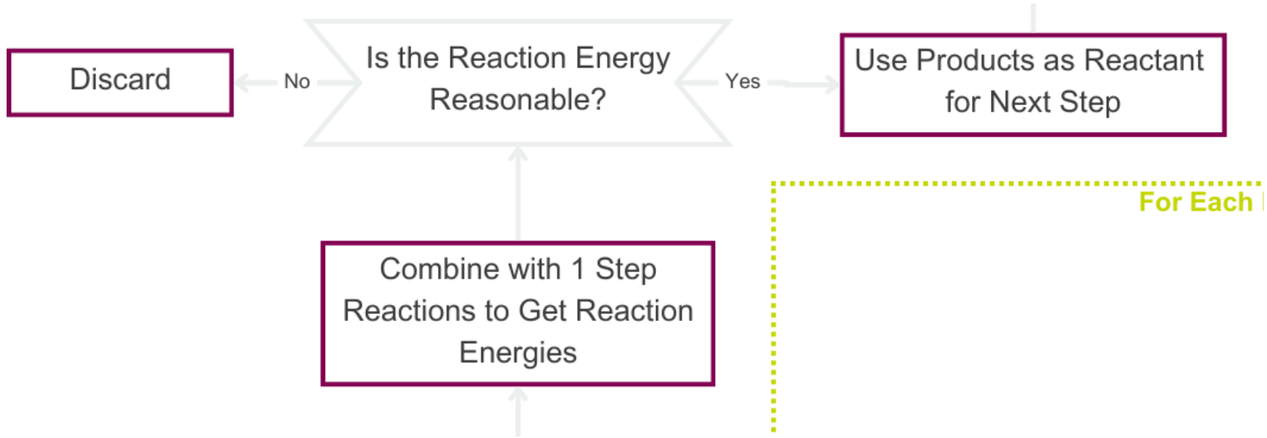
Calculating Energies

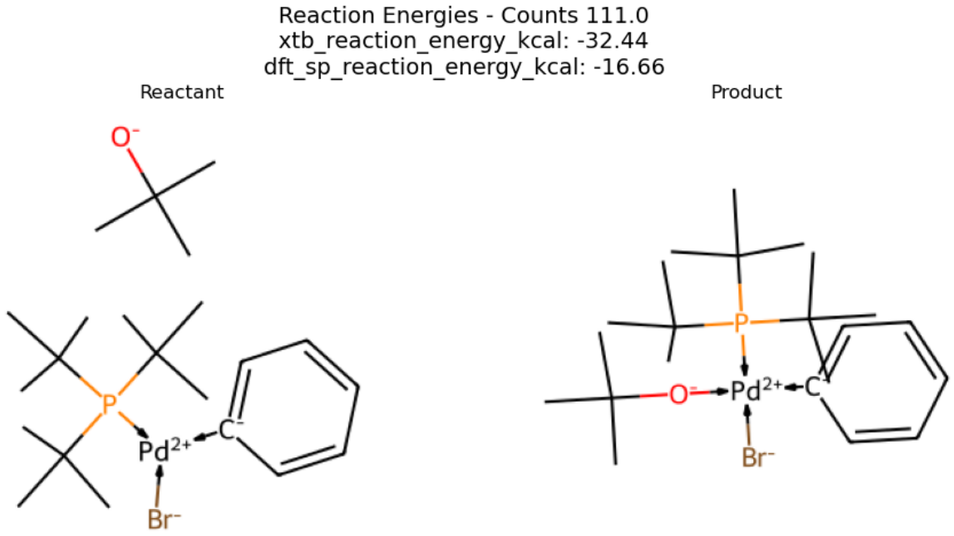
- GFN2 - xTB:
  - *Geometry Optimization*
  - *Initial Energies*
- DFT:
  - *Single point on xTB structures*
  - *Geometry optimization on the lowest energy conformer*



Reaction Energies

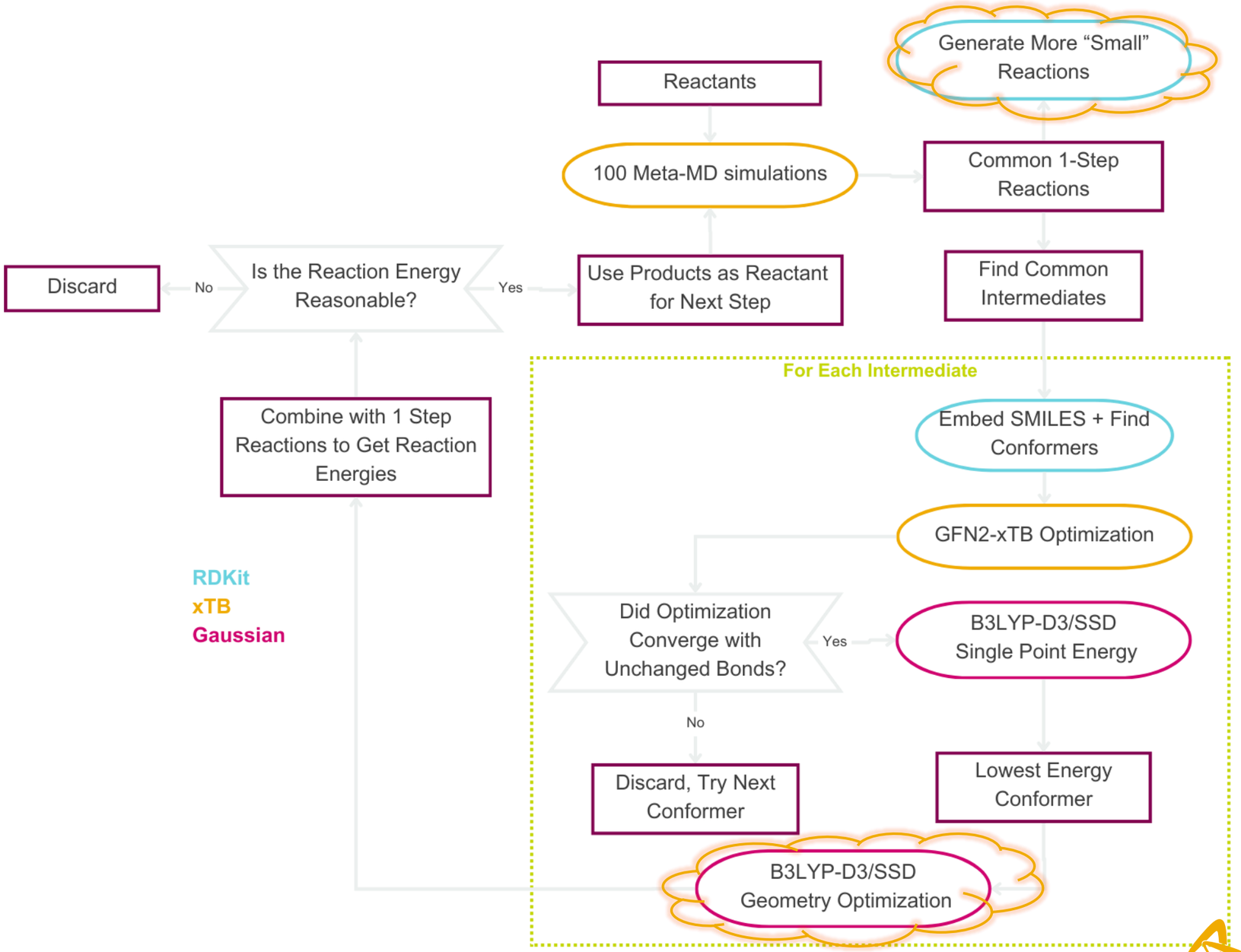
- Energy difference between reactant and product:
  - *Approximating ground state by choosing the lowest energy conformer*
- Mapping the energies from unique smiles to 1step to get reaction energies
- Choosing reaction energy interval to pick input for the next step
  - *Cycle starting over again*





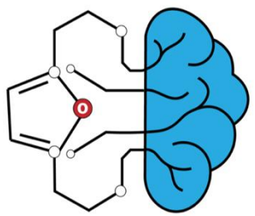
Next Steps

- Evaluate energies:
  - *Compare to article*
  - *Where to cut*
- Smaller reaction adding manual:
  - *Halide leaving*
  - *Tautomers*
- Transition States:
  - *Reaction barrier?*



UCPH

Thank you for listening!



UNIVERSITY OF COPENHAGEN

AstraZeneca