SAND2022-11536C

# *pynta:* An automated workflow code for reaction path exploration on surfaces

Matthew S. Johnson,[1] Maciej Gierada,[1] David H. Bross,[2] Eric D. Hermes,[1]
Christopher Blais,[3] C. Franklin Goldsmith,[4] Richard H. West,[3]
Khachik Sargsyan,[1] Habib N. Najm,[1] Judit Zádor[1]

[1] Combustion Research Facility, Sandia National Laboratories, Livermore, CA, USA
[2] Chemical Sciences and Engineering Division, Argonne National Laboratory, Lemont, IL, USA
[3] Department of Chemical Engineering, Northeastern University, Boston, MA, USA
[4] School of Engineering, Brown University, Providence, RI, USA

Open-Source Software for Kinetics, Chemical Networks, & Reactor Modeling

2022 Fall ACS Meeting

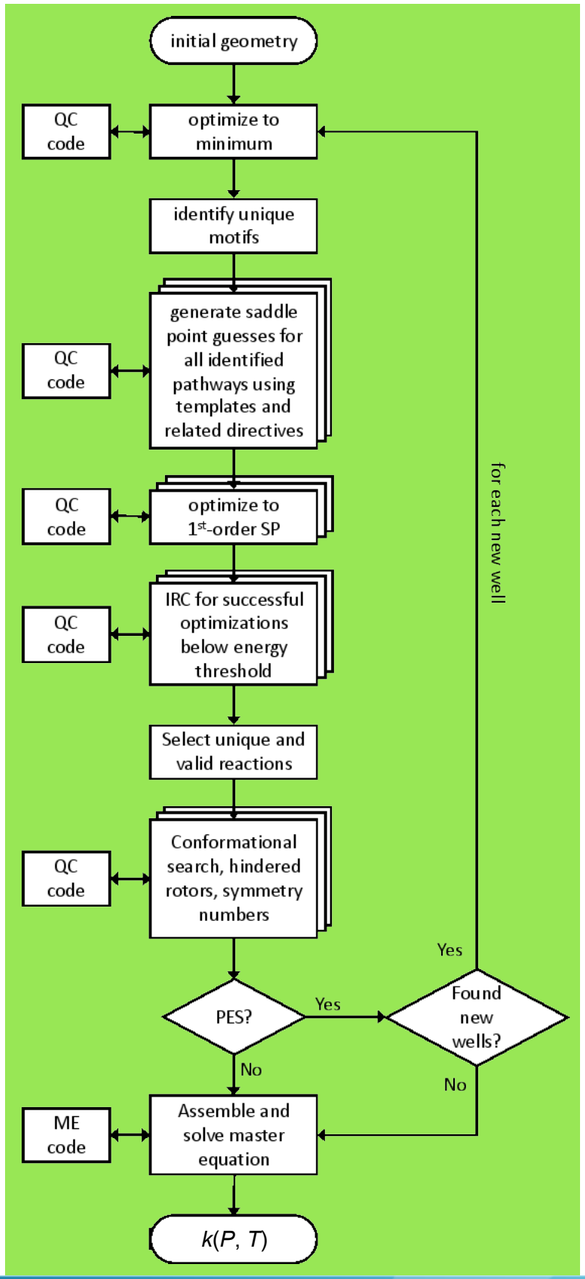
8/21 – 8/25, 2022

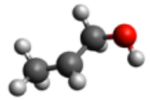
Funding: DOE BES CTC

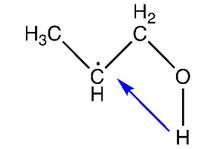

# Automation for gas-phase kinetics

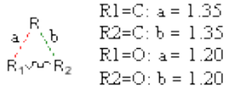
github.com/zadorlab/KinBot

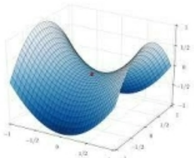
Van de Vijver, JZ, Computer Physics Communication, 2020, 248, 106947

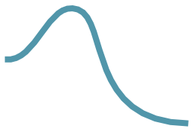


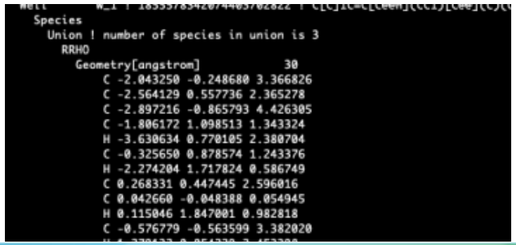



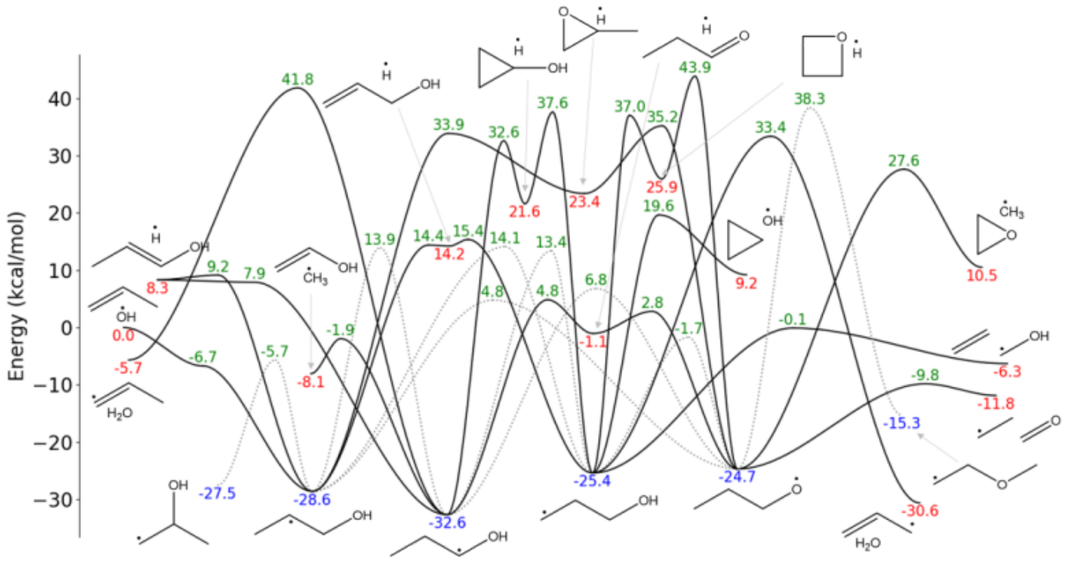


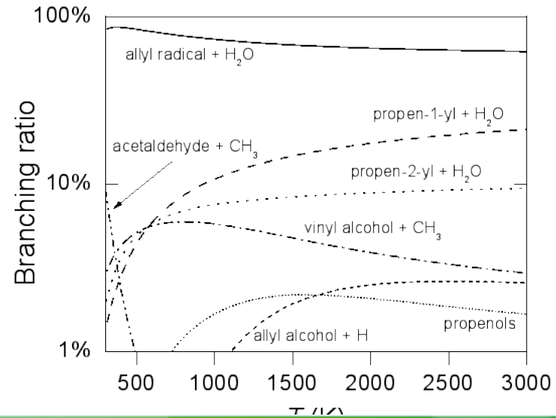



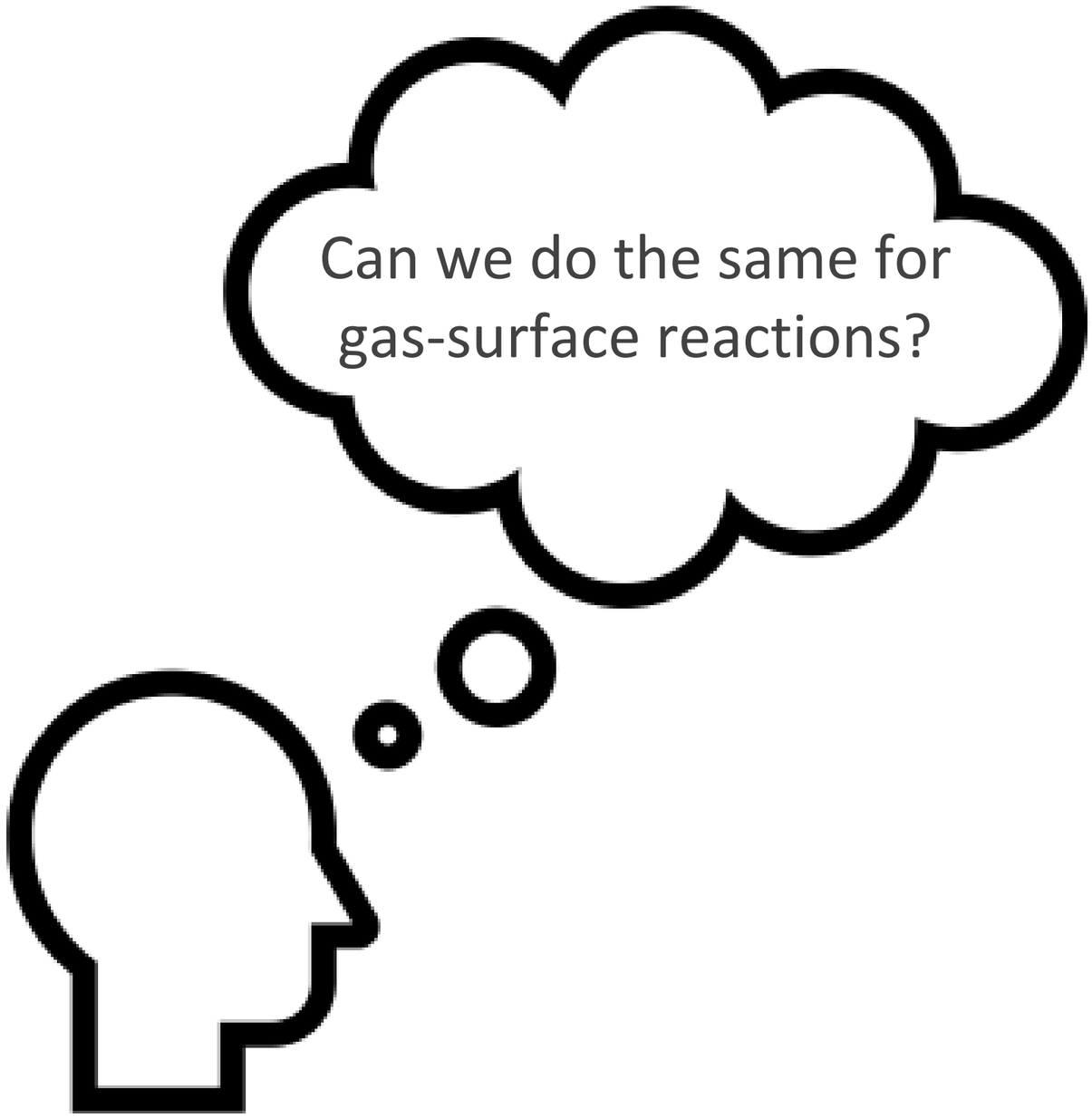







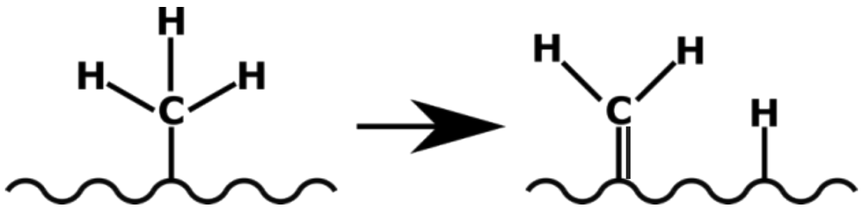



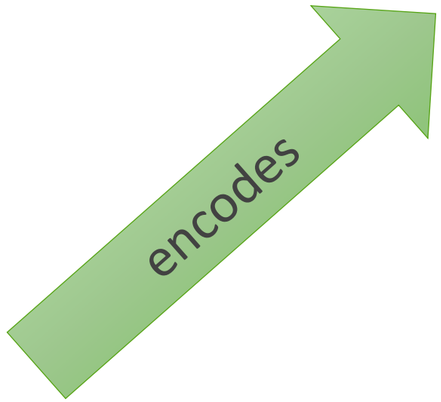

Can we do the same for gas-surface reactions?

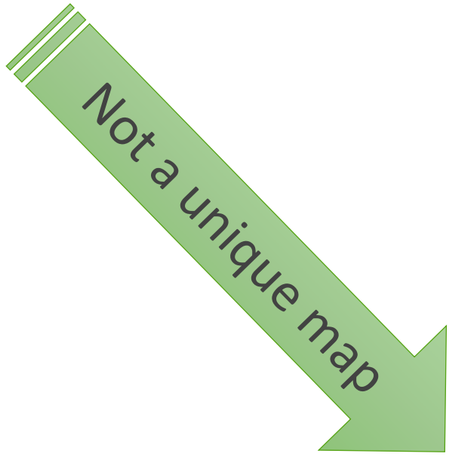
# From formal catalytic reactions to *ab initio* kinetics

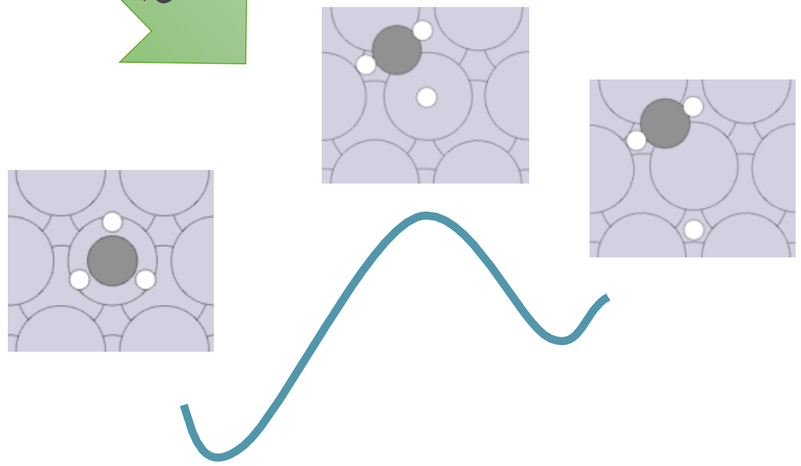
RMG reaction representation

```
- index: 0
  reaction: CH3X + X <=> CH2X + HX
  reaction_family: Surface_Abstraction
  reactant: |

    1 *1 C u0 p0 c0 {2,S} {3,S} {4,S} {5,S}
    2 *2 H u0 p0 c0 {1,S}
    3    H u0 p0 c0 {1,S}
    4    H u0 p0 c0 {1,S}
    5    X u0 p0 c0 {1,S}
    6 *3 X u0 p0 c0
  product: |

    1 *1 C u0 p0 c0 {3,S} {4,S} {5,D}
    2 *2 H u0 p0 c0 {6,S}
    3    H u0 p0 c0 {1,S}
    4    H u0 p0 c0 {1,S}
    5    X u0 p0 c0 {1,D}
    6 *3 X u0 p0 c0 {2,S}
```



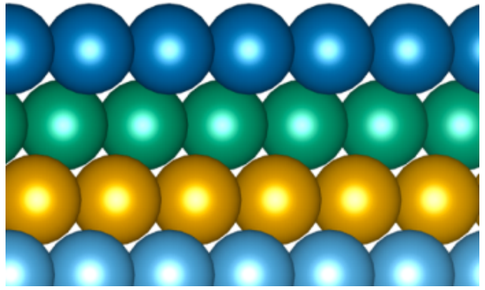



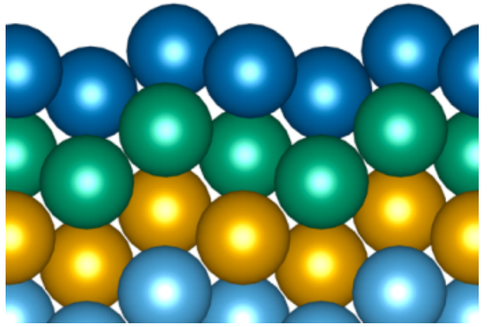
# There are many potential binding sites even for simple surfaces

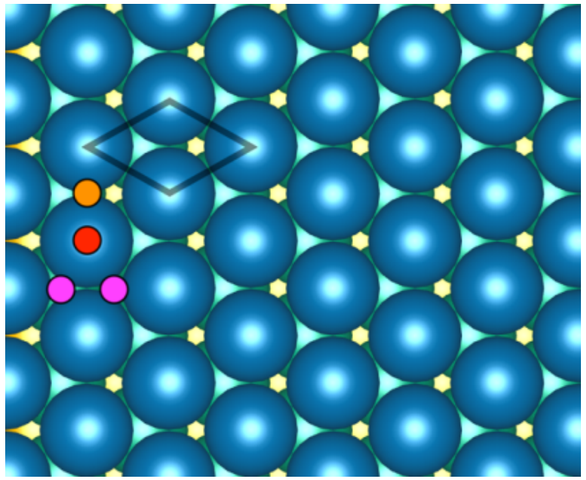
FCC(111)

FCC(211)

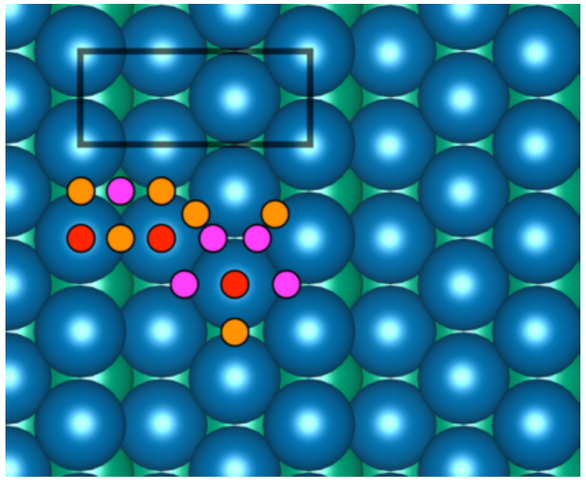


Side view





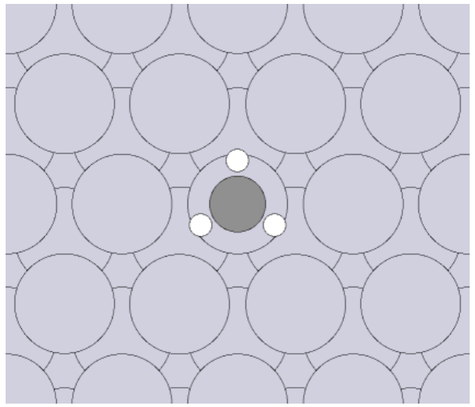
Top view



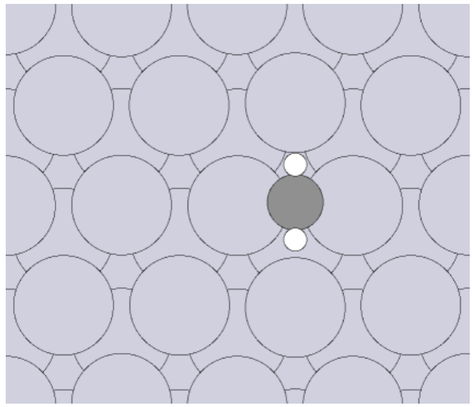
**...and there are even more for high-index metal facets or nanoparticles.**

# There are only a few sites that are minima on Pt(111) for these species

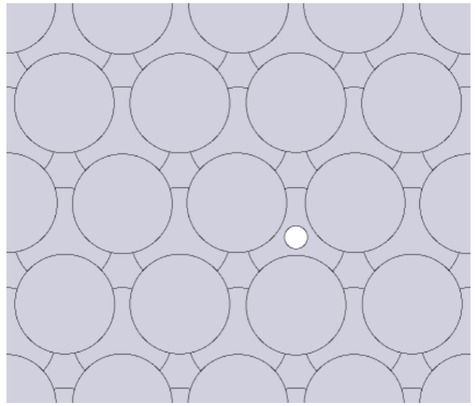
$CH_3$

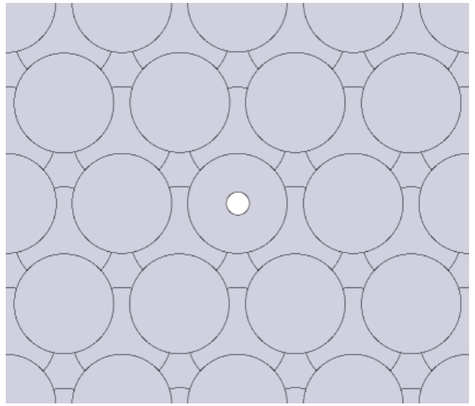


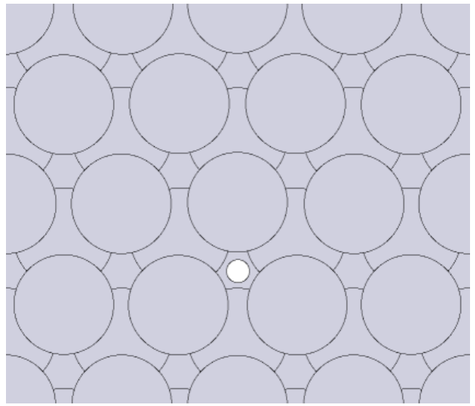
$CH_2$



H

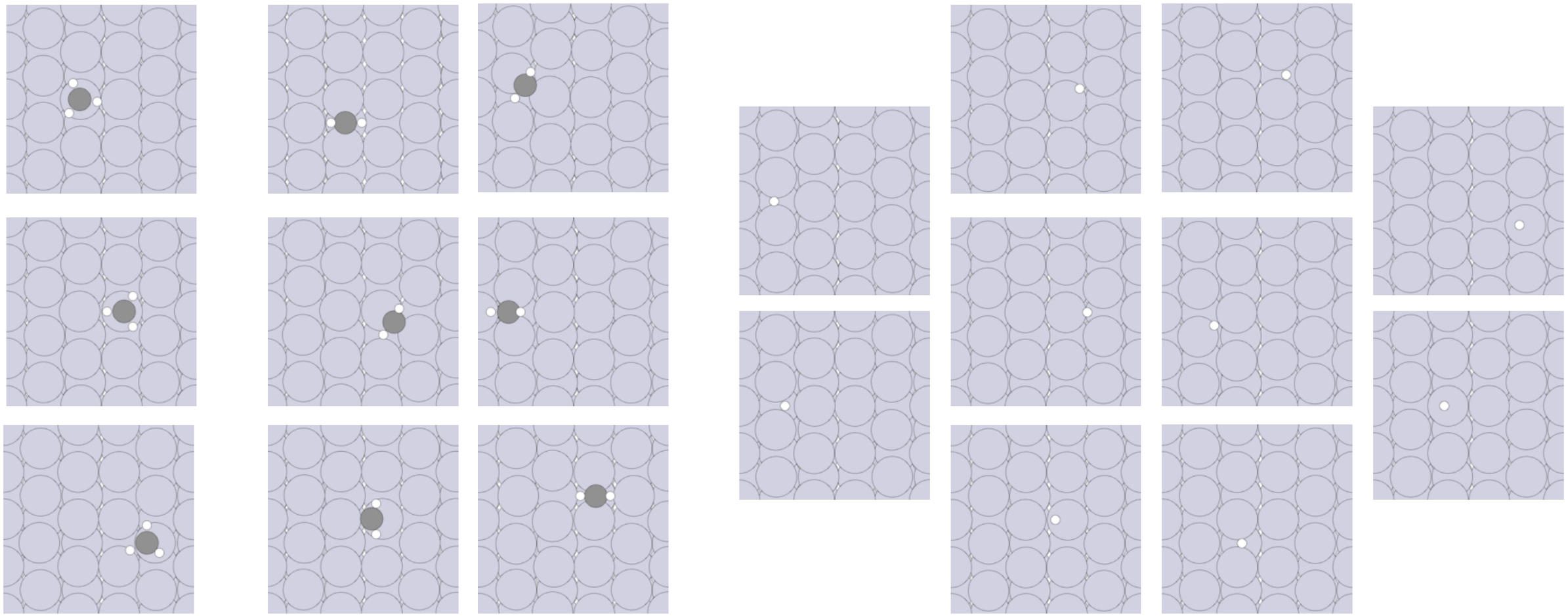






# But there are a lot more already on Pt(211)

$CH_3$ $CH_2$ H

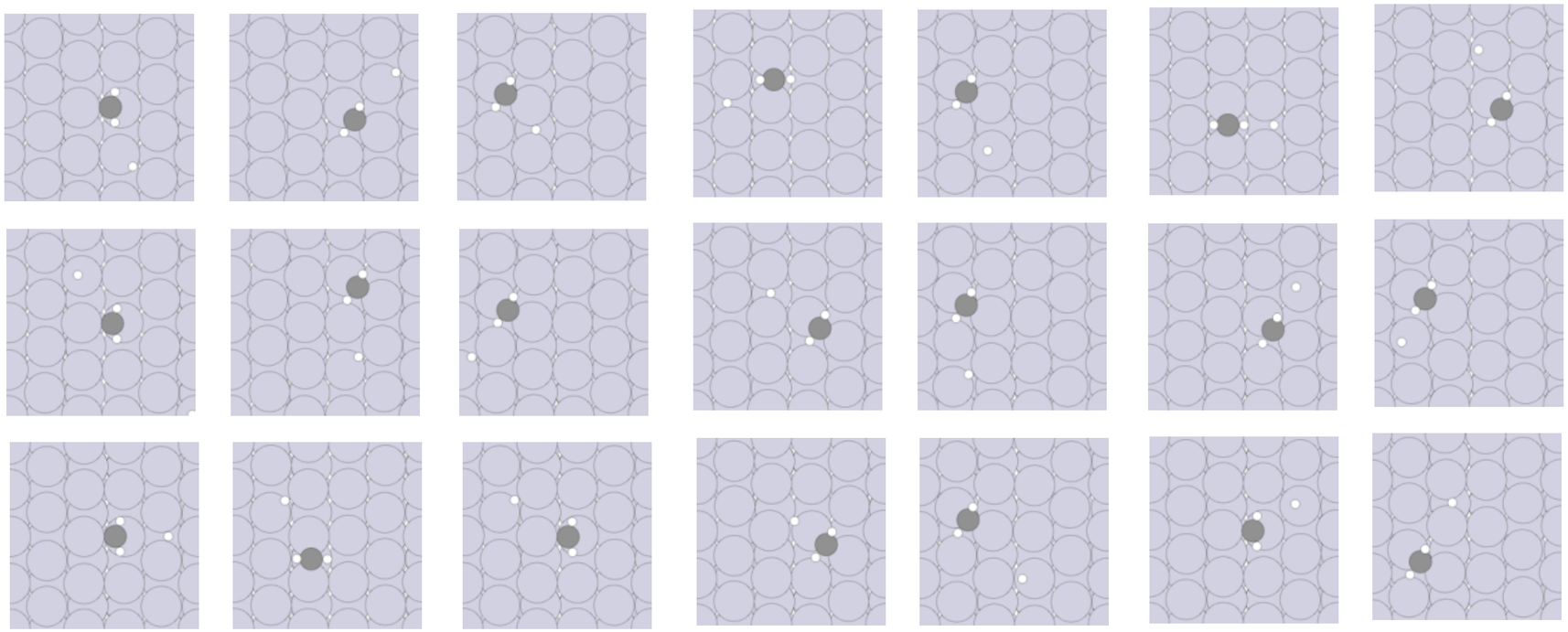


Thermochemistry can be strongly dependent on the choice of site

**Pt(111):** $\Delta_r E$: 0.267 +/- 0.053 eV

**Pt(211):** $\Delta_r E$: 0.064 +/- 0.415 eV

# $CH_2$ + H co-adsorbed minima on Pt(211)



…There are **73** total minima!

It is difficult or impossible to know a priori which minima are directly connected to the reaction of interest, and which involve diffusion of adsorbates in addition to reaction.

➔Can we still do a systematic search at a cost lower than a full NEB search?

We started developing ***pynta*** https://github.com/zadorlab/pynta
(“pinta” means surface in Finnish)

# Initial workflow for dissociation/association reactions

1. Prepare surface → Saved and reused
    - Size
    - Symbol
    - Lattice constant
    - Vacuum
    - Type
    - Level of theory
2. Place reactant and product adsorbate(s) extracted from RMG notation on all unique sites

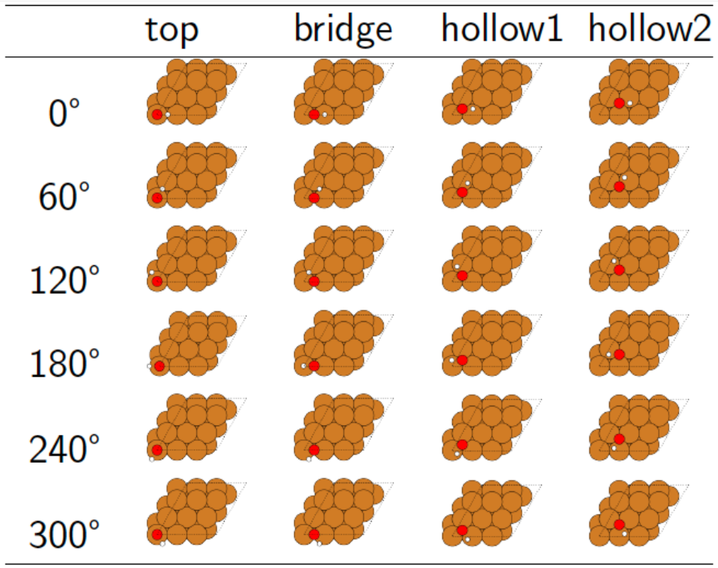
    ➔Optimize, filter

    ➔Calculate vibrational frequencies

    ➔This step provides thermo information
        - Lowest energy structures
        - Aids more accurate partition function calculations

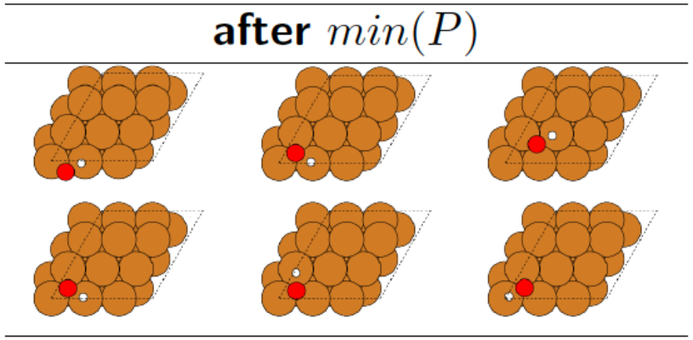
# Initial workflow for dissociation/association reactions

3. Create saddle point guesses
   - Place unimolecular species (AB) on all sites and create rotamers
   - Filter




   - Stretch bond to 1.4x and apply penalty function, surface is not relaxed

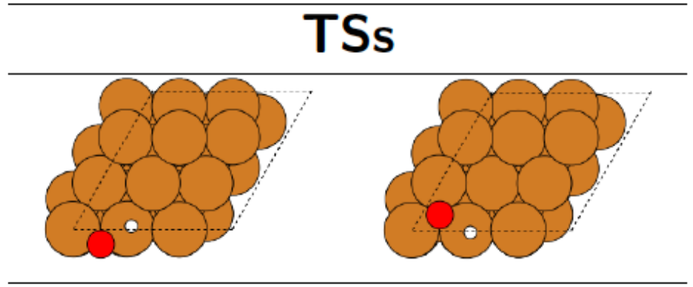
$$P = (r_{\mathrm{A-X}} - \overline{r_{\mathrm{A-X}}})^2 + (r_{\mathrm{B-X}} - \overline{r_{\mathrm{B-X}}})^2 + E_{\mathrm{xTB}}$$




   - Filter

# Initial workflow for dissociation/association reactions

4. Optimize to saddle points with all constraints lifted at DFT level
   - Geometry optimization is done with **Sella**, our own optimizer (sella means saddle in Latin)

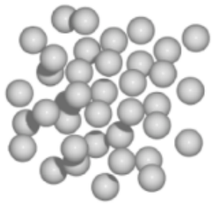



https://github.com/zadorlab/sella

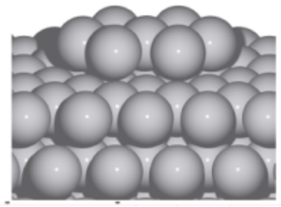
E. D. Hermes, K. Sargsyan, H. N. Najm, JZ: JCTC. 2019, 15 6536-6549.
E. D. Hermes, K. Sargsyan, H. N. Najm, JZ: JCP, 2021, 155 094105.
E. D. Hermes, K. Sargsyan, H. N. Najm, JZ: JCTC, 2022, submitted.

Sella has superior performance for saddle points optimizations for solids and molecular systems, and has featureful optimizations with constraints, and can carry out IRC as well.

https://www.optbench.org

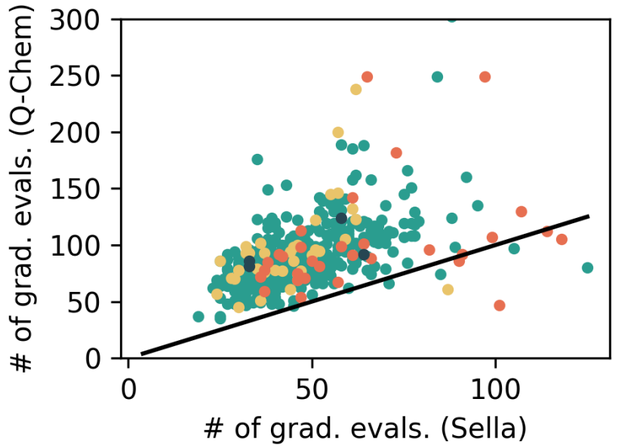


| Algorithm | <FCs> | min(FCs) | median(FCs) | max(FCs) | nfailed |
|---|---|---|---|---|---|
| Sella | 81 | 37 | 75 | 166 | 0 |
| optim_ffd | 145 | 57 | 119 | 565 | 0 |
| pele_ffd | 192 | 59 | 150 | 1488 | 0 |
| eon-lanczos-improved[1] | 237 | 65 | 181 | 1898 | 0 |
| eon-lanczos[1] | 321 | 83 | 220 | 8857 | 0 |
| eon-dimer[1] | 528 | 92 | 388 | 3581 | 0 |
| Show/Hide Additional Entries... | | | | | |



| Algorithm | <FCs> | min(FCs) | median(FCs) | max(FCs) | nfailed |
|---|---|---|---|---|---|
| Sella | 49 | 32 | 45 | 98 | 0 |
| optim_ffd | 71 | 43 | 70 | 143 | 0 |
| pele | 88 | 52 | 78 | 198 | 0 |
| optim_cd | 90 | 52 | 86 | 199 | 0 |
| eon-lanczos-improved[1] | 106 | 71 | 100 | 163 | 0 |
| eon-dimer[1] | 116 | 83 | 114 | 160 | 0 |
| eon-lanczos[1] | 143 | 106 | 139 | 292 | 0 |
| Show/Hide Additional Entries... | | | | | |

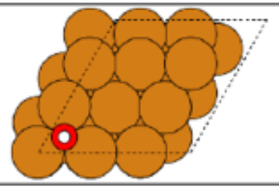
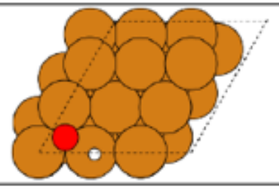
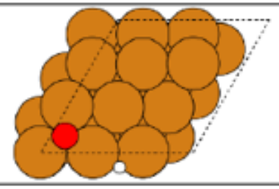
Example comparison from our own molecular saddle benchmark



# Initial workflow for dissociation/association reactions

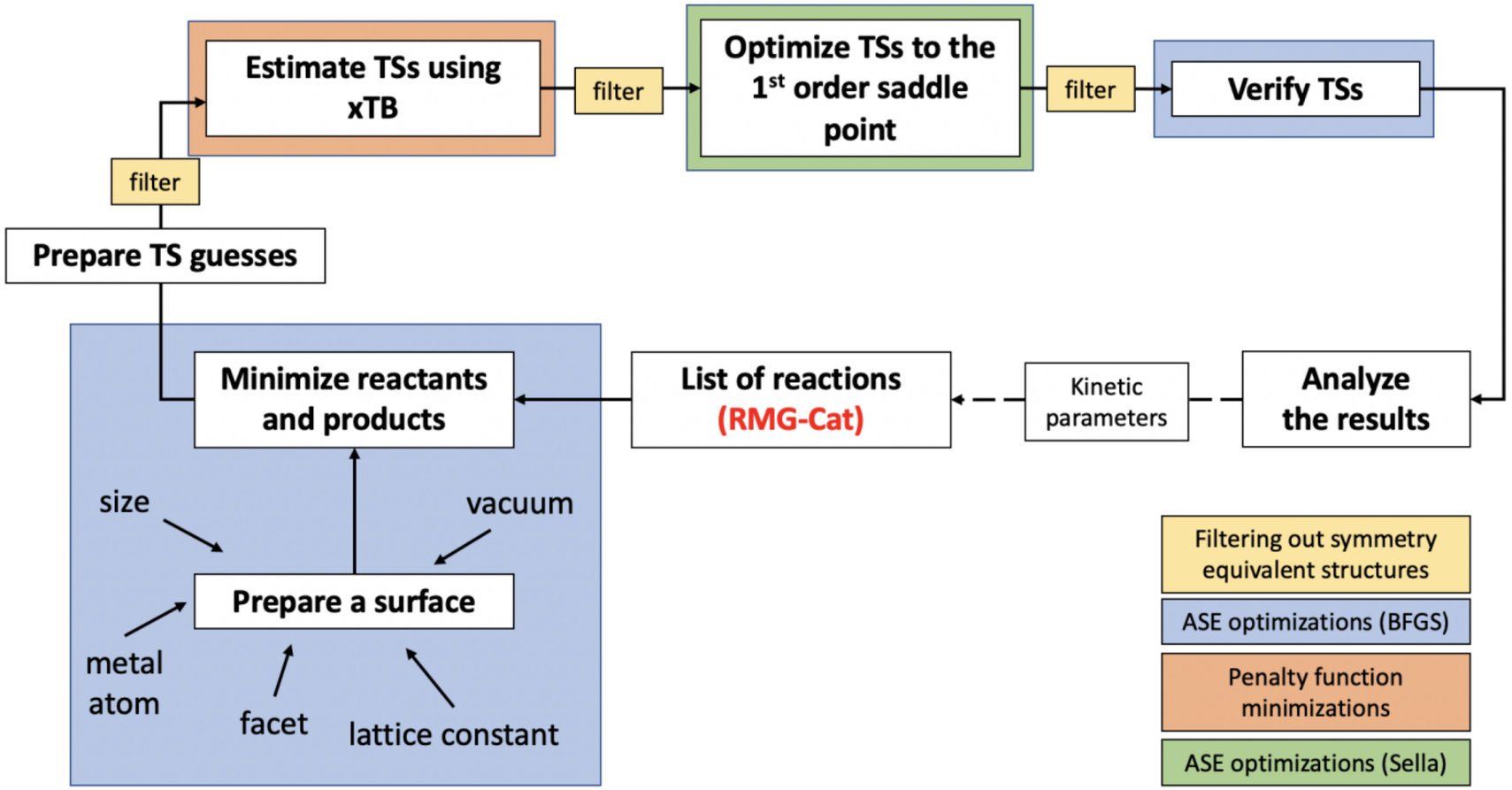
5. Verification
   - IRC
   - Frequency calculations
6. Analysis
   - Reaction barrier
   - Reaction enthalpy
   - Geometry

$$OH^* + {}^* \rightleftharpoons O^* + H^*$$

| # | reactants | TS | products | $\Delta E^r$ / kJ · mol$^{-1}$ | $E^{\ddagger}$ / kJ · mol$^{-1}$ | $r^{\ddagger}_{C-H}$ / Å |
|---|---|---|---|---|---|---|
| 00 |  |  |  | 69.155 | 185.321 | 1.71 |

7. Passing data to rate coefficient calculator

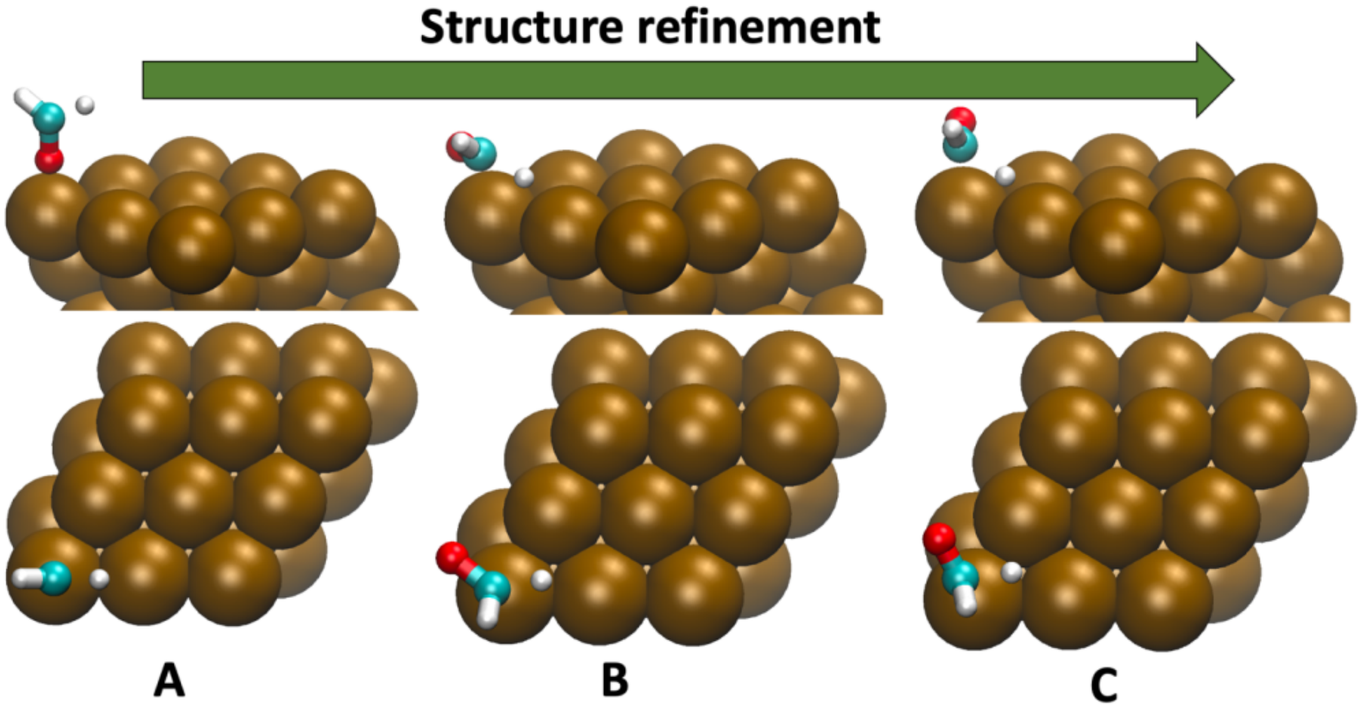
# Summary and computational aspects




Important dependencies

- Balsam
- ASE
- Sella
- xTB
- Catkit

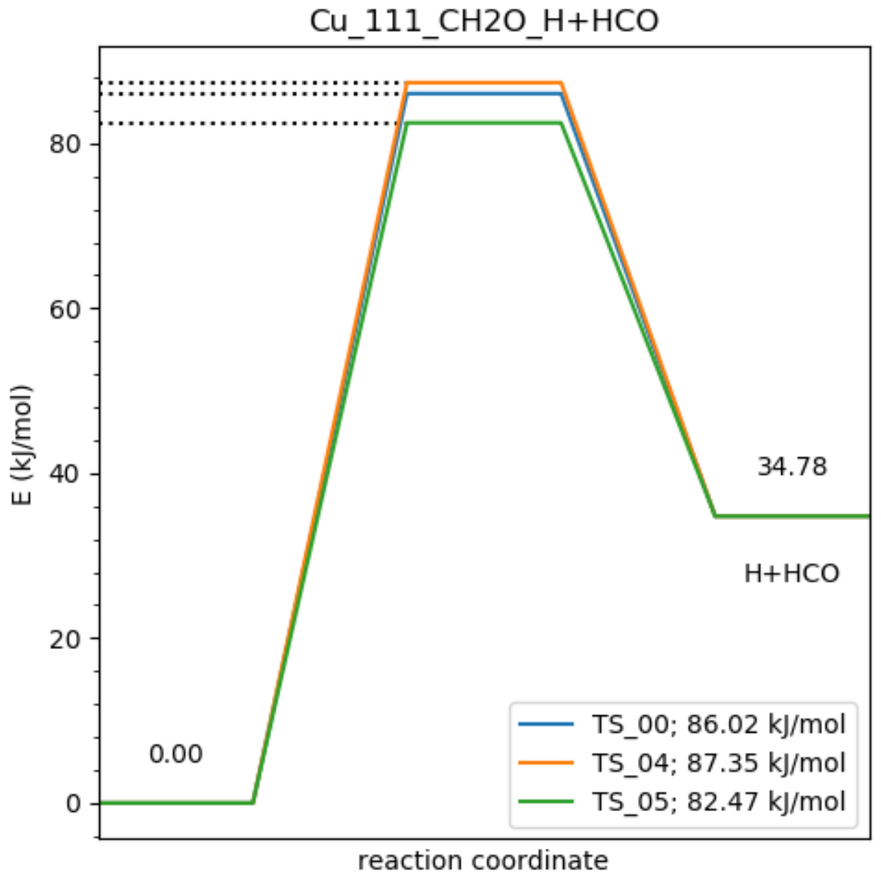
# $CH_2O$ → CHO + H on Cu(111)




**“raw” guess**
4 sites
292 initial structures

**penalty fn optimization**
42 unique starting geoms
10 final ones

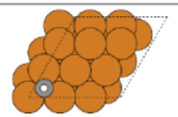
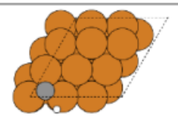
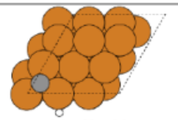
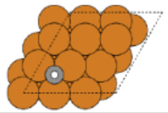
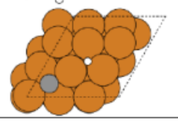
**final saddle point**
3 unique saddles



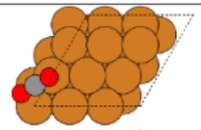
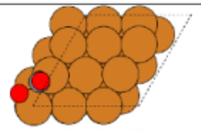
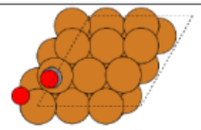
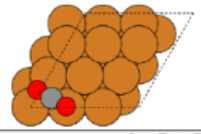
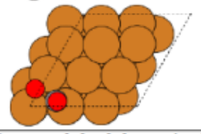
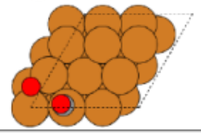

PBE functional results
Literature lists only one saddle

# Further examples

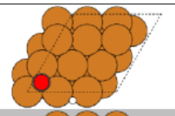
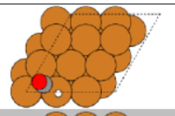
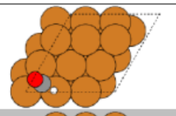
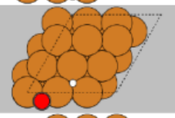
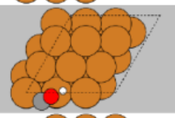
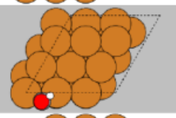
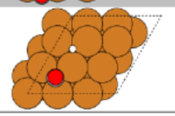
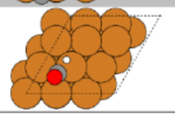
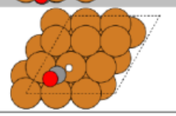
$CH^* + {}^* \rightleftharpoons C^* + H^*$

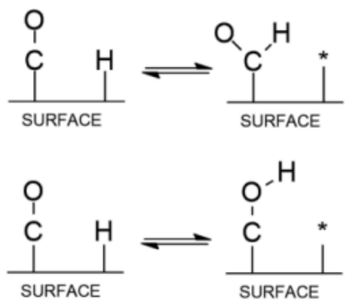
| # | reactants | TS | products | $\Delta E^r$ (kJ · mol$^{-1}$) | $E^\ddagger$ (kJ · mol$^{-1}$) | $r^\ddagger_{C-H}$ (Å) |
|---|---|---|---|---|---|---|
| 00 |  |  |  | 132.836 | 183.297 | 1.82 |
| 01 |  |  |  | 117.379 | 197.932 | 2.00 |

$CO^* + O^* \rightleftharpoons CO_2^* + {}^*$

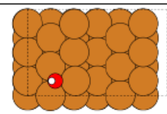
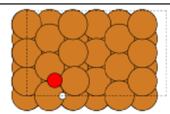
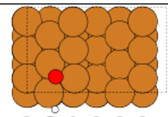
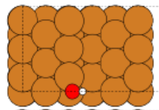
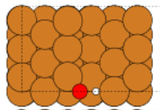
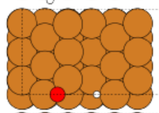
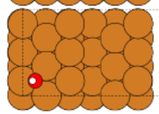
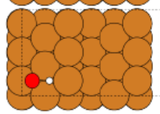
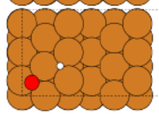
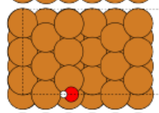
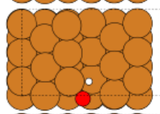
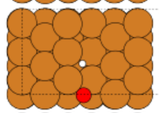
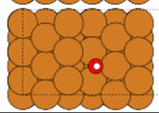
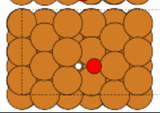
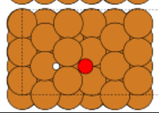
| | reactants | TS | products | $\Delta E^r$ | $E^\ddagger$ | $r^\ddagger_{C-O}$ |
|---|---|---|---|---|---|---|
| 00 |  |  |  | -111.682 | 50.449 | 1.77 |
| 01 |  |  |  | -118.958 | 46.756 | 1.82 |

$CO^* + H^* \rightleftharpoons HCO^* + {}^*$

| # | reactants | TS | products | $\Delta E^r$ (kJ · mol$^{-1}$) | $E^\ddagger$ (kJ · mol$^{-1}$) | $r^\ddagger_{X-H}$ (Å) |
|---|---|---|---|---|---|---|
| 00 |  |  |  | 93.809 | 103.403 | 1.52 |
| 01 |  |  |  | 123.530 | 223.419 | 1.28 |
| 02 |  |  |  | 86.842 | 96.120 | 1.41 |




$OH^* + {}^* \rightleftharpoons H^* + O^*$

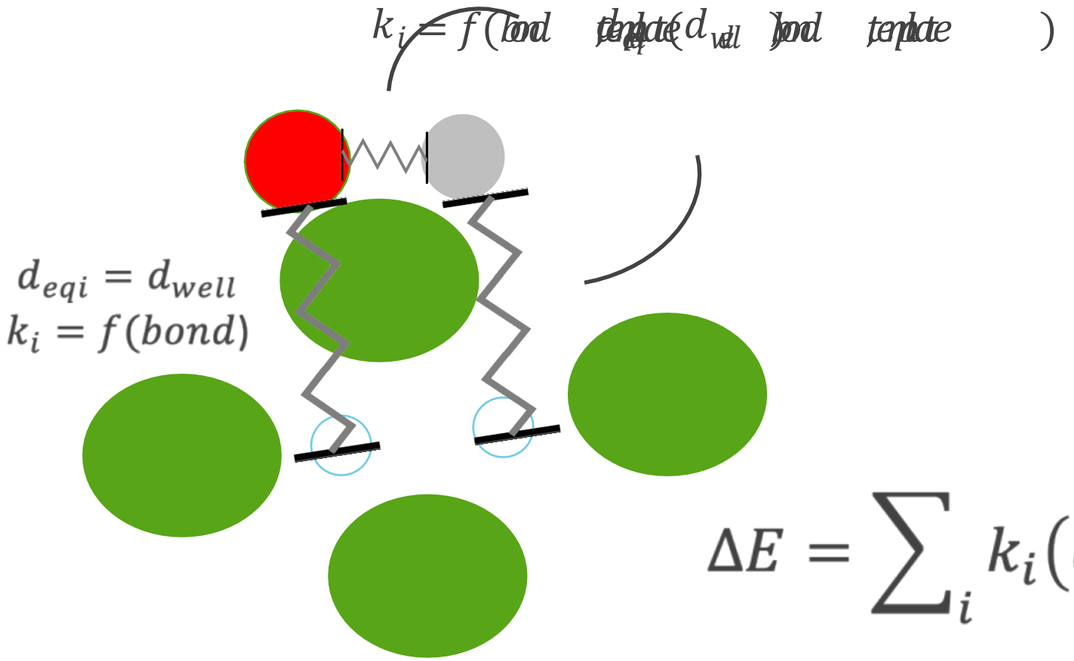
| # | reactants | TS | products | $\Delta E^r$ (kJ · mol$^{-1}$) | $E^\ddagger$ (kJ · mol$^{-1}$) | $r^\ddagger_{X-H}$ (Å) |
|---|---|---|---|---|---|---|
| 03 |  |  |  | 69.598 | 157.963 | 1.57 |
| 05 |  |  |  | 95.090 | 148.133 | 1.69 |
| 07 |  |  |  | 63.890 | 150.247 | 1.55 |
| 08 |  |  |  | 114.593 | 200.130 | 1.64 |
| 11 |  |  |  | 23.954 | 107.021 | 1.53 |

# In Progress and Future Plans

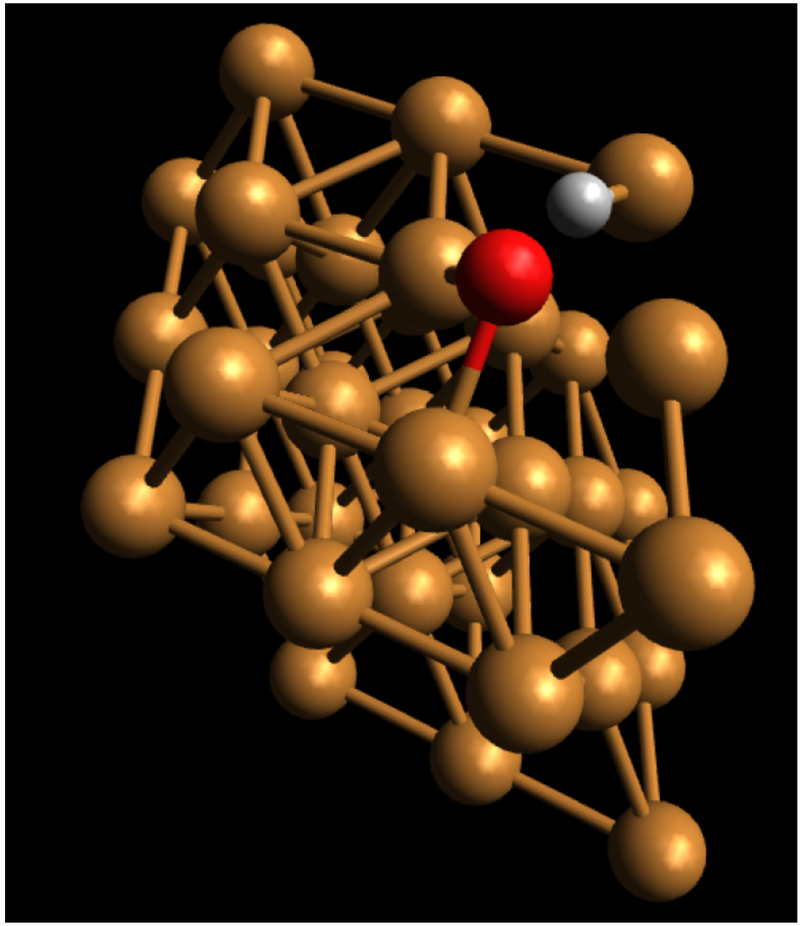
- Integrate Pynta with the Fireworks workflow management software to
  - Operate efficiently on larger clusters
  - Improve restart functionality
  - Enable automatic workflow debugging
  - Make it easy to adapt the workflow as needed
- Integrate IRC calculations using Sella for TS verification
- Generalize TS estimation to most reaction templates of interest
  - Analyze reaction templates using RMG and the optimized reactant and product adsorbates using ASE
  - Use this information to generate a modified xTB potential that will optimize starting structures to good transition state guesses for the reaction of interest
- Integrate database for storing useful generated information from Pynta
  - Enable reuse of prior calculations
  - Make results accessible to scientific community

# Harmonically Forced TS Searching

OHX <=> OX + HX

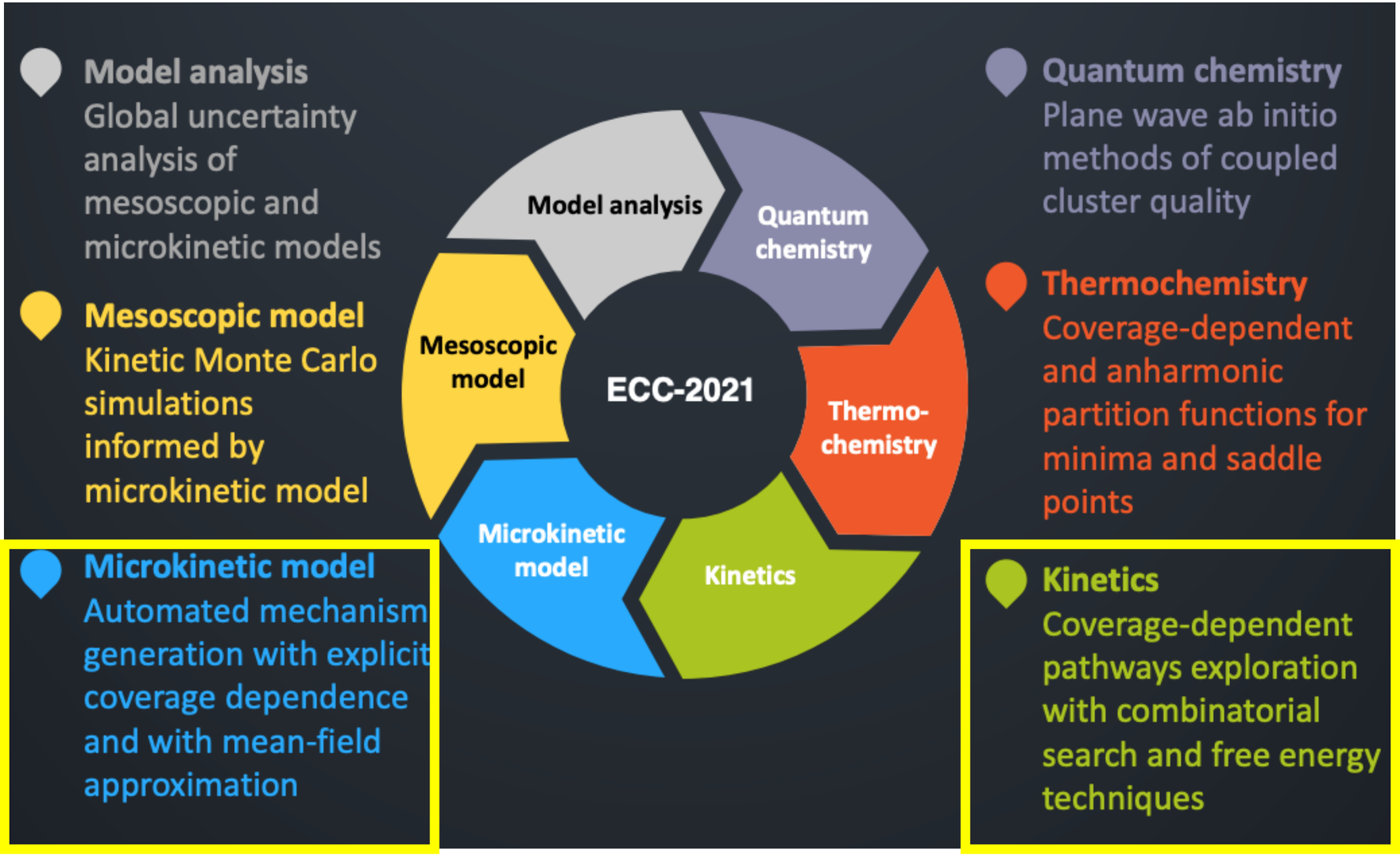



$$\Delta E = \sum_i k_i \left(d_i - d_{eqi}\right)^2$$



Where $d_{\mathrm{eq}}$ is the equilibrium bond length, $k_i$ is the spring constant of the bond, $d_{\mathrm{well}}$ is the bond length in the associated adsorbed structure and $d_i$ is the bond length.

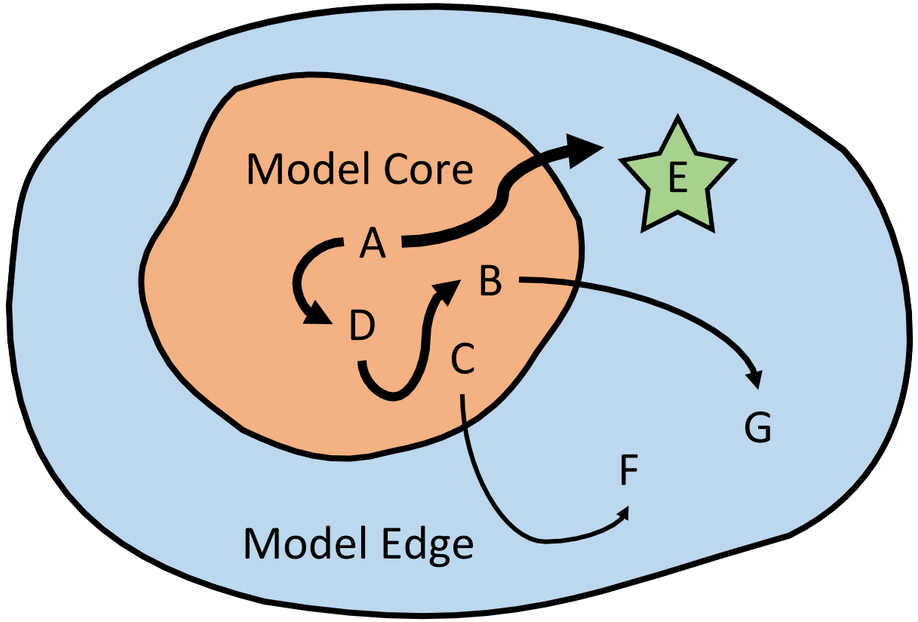
# pynta is part of the Exascale Catalytic Chemistry Project

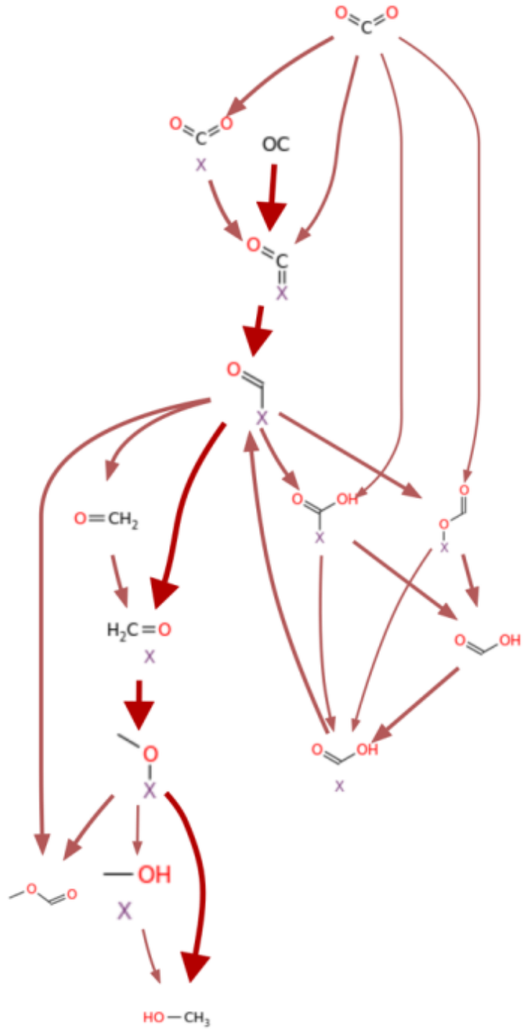



https://www.ecc-project.org

# RMG can generate microkinetic models for catalytic processes

https://github.com/ReactionMechanismGenerator/RMG-Py



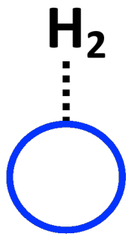

## Methanol synthesis on Cu(111)

- Well studied mechanism
- Data exists for both experimental (Graaf) and theoretical work (Grabow)
- Good proof of concept mechanism for RMG


L. C. Grabow, ***ACS Catal.*** 2011, 1, 4, 365–384.
G. H. Graaf, ***Chem. Eng. Sci.*** 1988, 43, 3185.

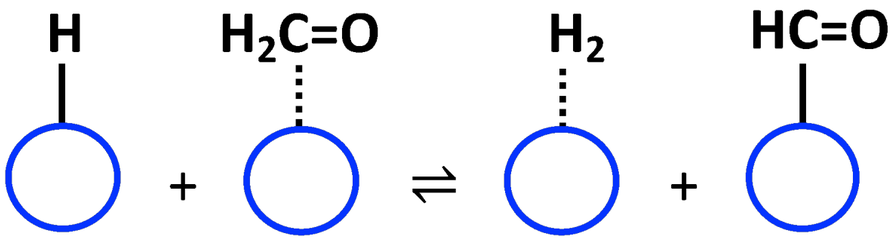
# van der Waals (vdW) bonded $H_2$ was choking the surface

Literature data (Álvarez-Falcón) was incorporated for the binding of $H_2$ on Cu(111).
RMG was estimating the thermo from Pt(111).
Unexpectedly, we were underestimating the binding energy.




DFT BE on Pt(111): -0.054 eV.
DFT BE on Cu(111): -0.08 eV.

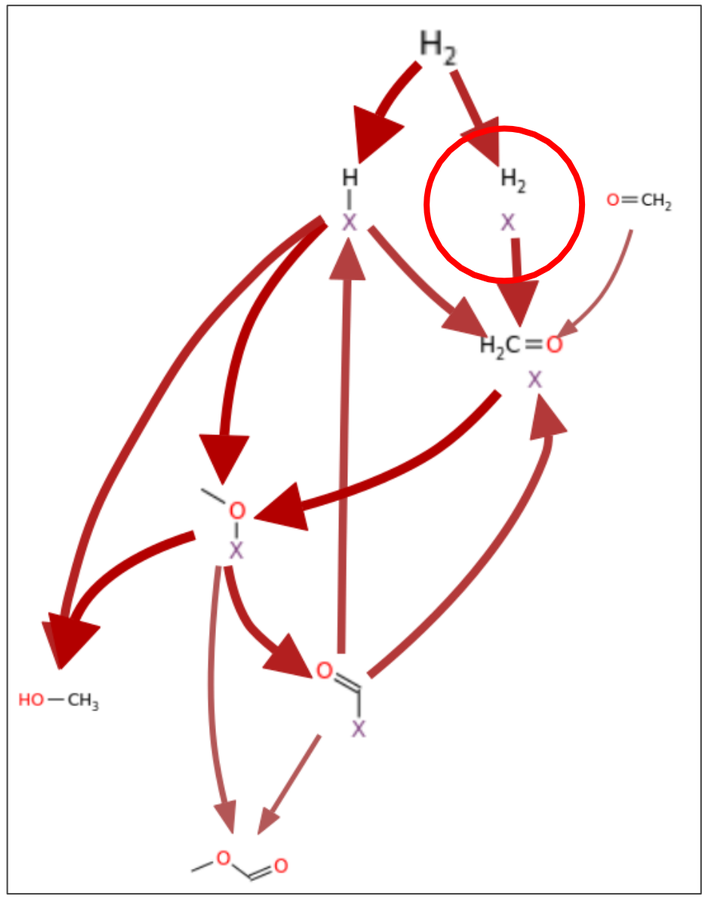
Sensitivity analysis revealed that the most sensitive reaction in the mechanism involved this $H_2$ species:



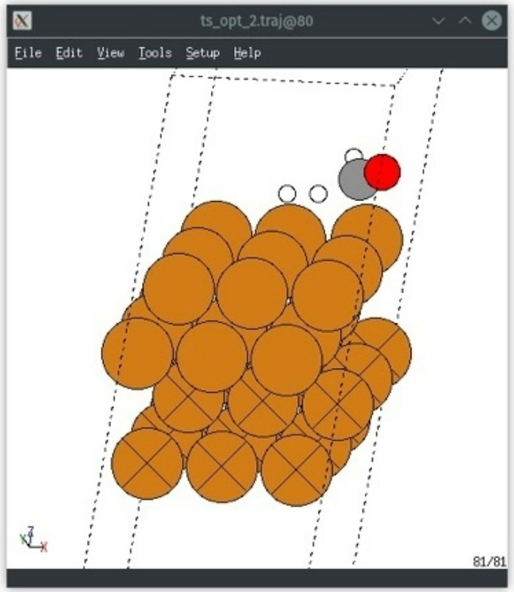

$k_f = 4.0 \times 10^5\ s^{-1}$ at 298K

L. Álvarez-Falcón, ***SurfSci,*** 2016, 646, 221-229.

# A saddle point search in pynta revealed the reaction was possible but very slow.

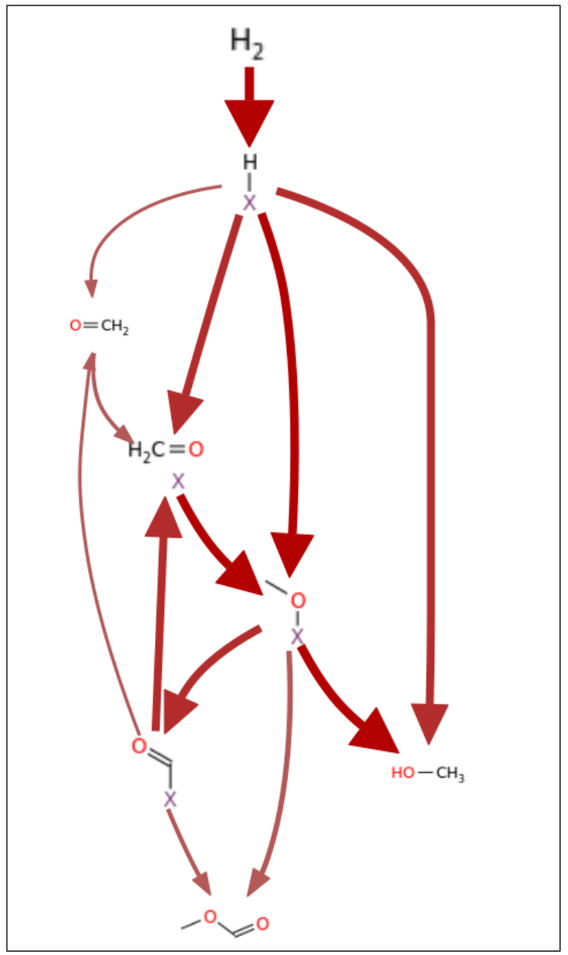



Model With Estimated Kinetics




$k_f = 4.0 \times 10^{5}\ s^{-1}$ at 298K

$k_f = 6.5 \times 10^{-14}\ s^{-1}$ at 298K




Model With DFT calculated Kinetics

Further updates of the model:

**Reaction enthalpies (kJ/mol)**

Cu_111_O+HCO_OH+CO -143.34
Cu_111_HCOOH_OH+HCO 76.33
Cu_111_H+HCO_CH2O -9.88
Cu_111_CH3OH_H+CH3O -12.01
Cu_111_CO2_O+CO 38.83
Cu_111_H+CH2O_CH3O -90.99
Cu_111_H+CO2_COOH 33.79
Cu_111_HCO+CH2O_CO+CH3O -184.8
Cu_111_CO2+HCO_CO+COOH -60.03
Cu_111_H+CO_HCO 93.81
Cu_111_CH2O_H+HCO 9.88

# Conclusions

- We are developing an open-source code, pynta, to explore surface-reactions automatically.
- The tool is designed to work with RMG, an automated reaction mechanism generator, which provides the reactions in the form of adjacency matrices.
- We demonstrated the viability of our approach in pynta through several examples, and outlined some of the immediate improvements that are currently under development.
- We provided an example in which pynta has improved a microkinetic model generated by RMG.