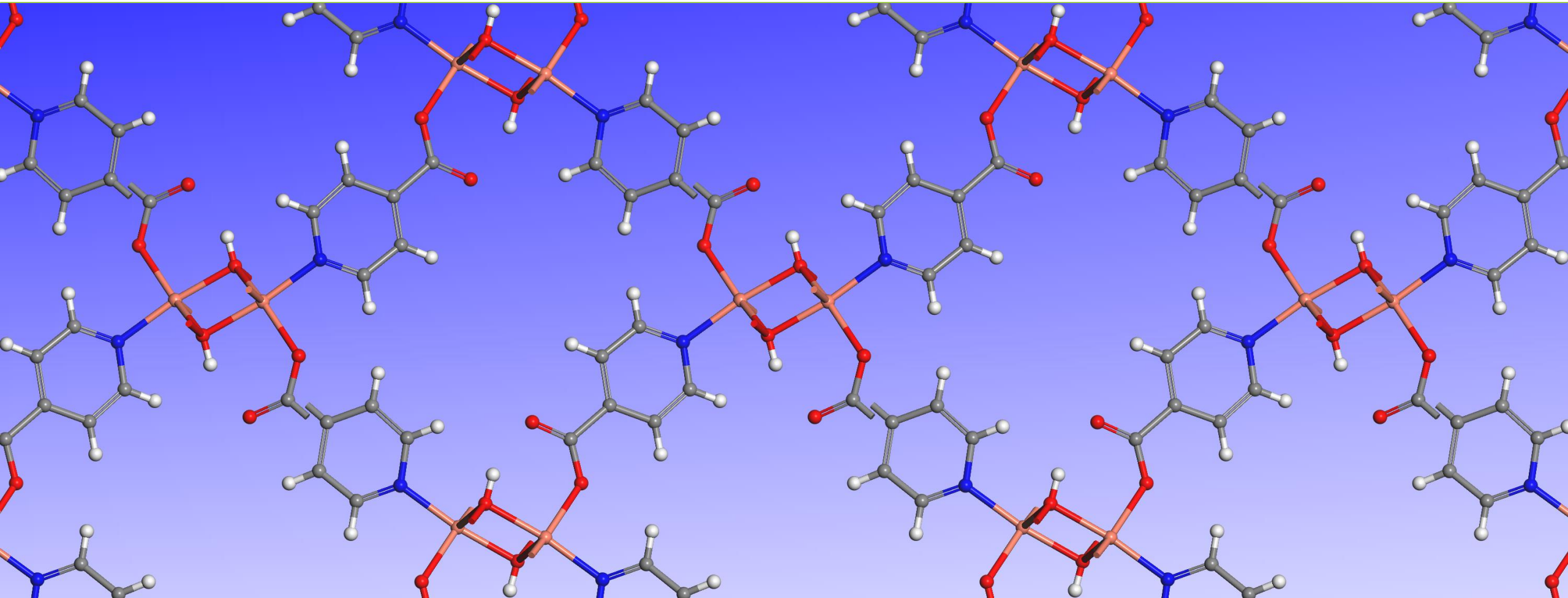
Machine Learned Force Field Modeling of Metal Organic Frameworks for CO_2 Direct Air Capture

John Findley, Samir Budhathoki, Jan Steckel

NATIONAL ENERGY TECHNOLOGY LABORATORY

ACS Meeting, August 2024



Acknowledgements

Co-authors:

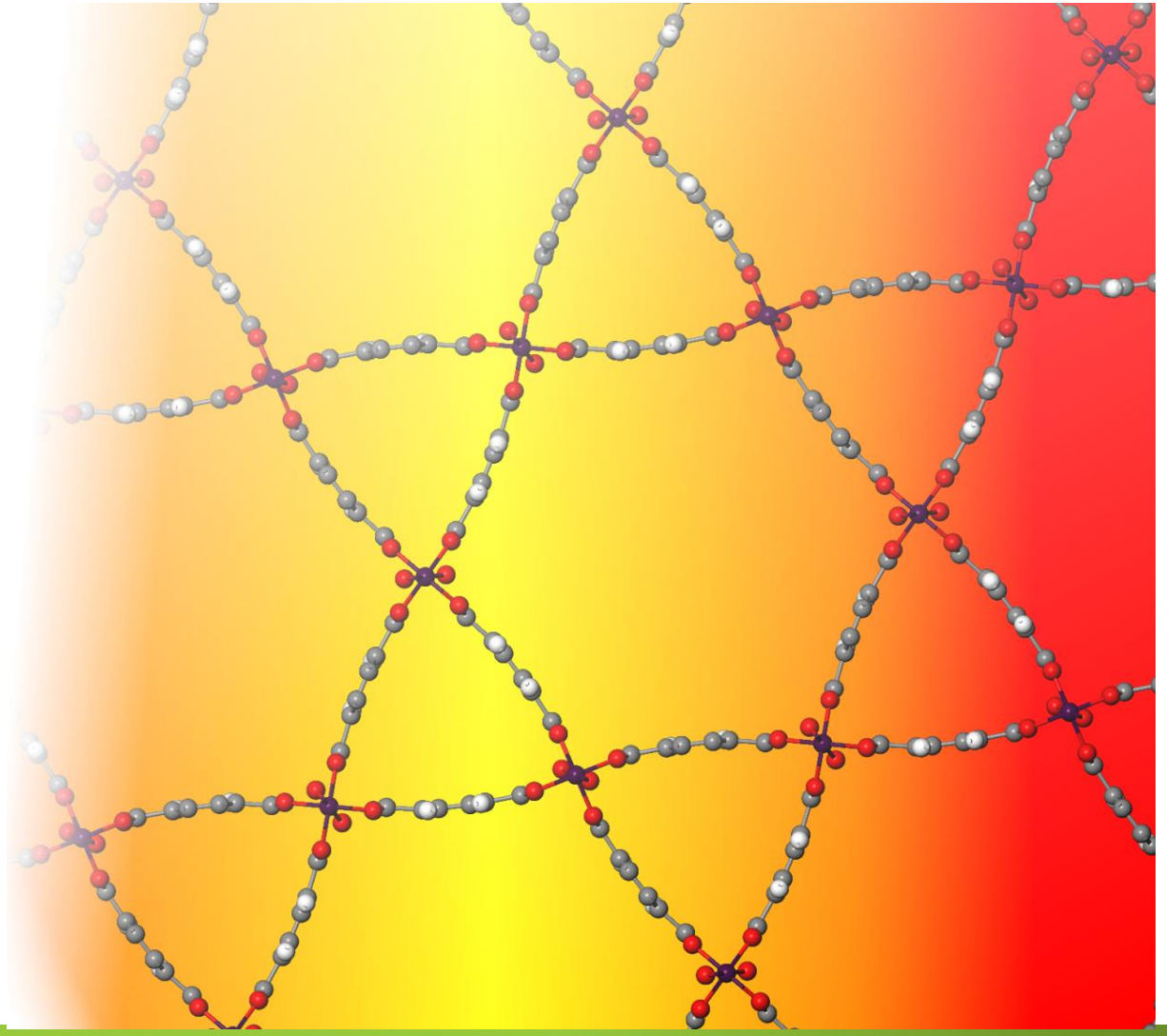
- Jack Findley – NETL
- Samir Budhathoki – formerly at NETL, currently Braskem

Summer Students:

- Tiernan Baucom, Ugrad, currently at University of Michigan
- Clare Martin, Ugrad, currently at University of Michigan
- Laine Roper, Ugrad, currently at Harvard

Collaborators:

- Christopher E. Wilmer, U of Pittsburgh
- Dan Sorescu - NETL



Outline and Disclaimer

Outline: Metal Organic Framework Modeling using three methods:

1. Classical (Rigid) Force Field: Universal Force Field (UFF)
2. Classical Flexible Force Field: QuickFF
3. Machine Learned Force Field: FitSNAP

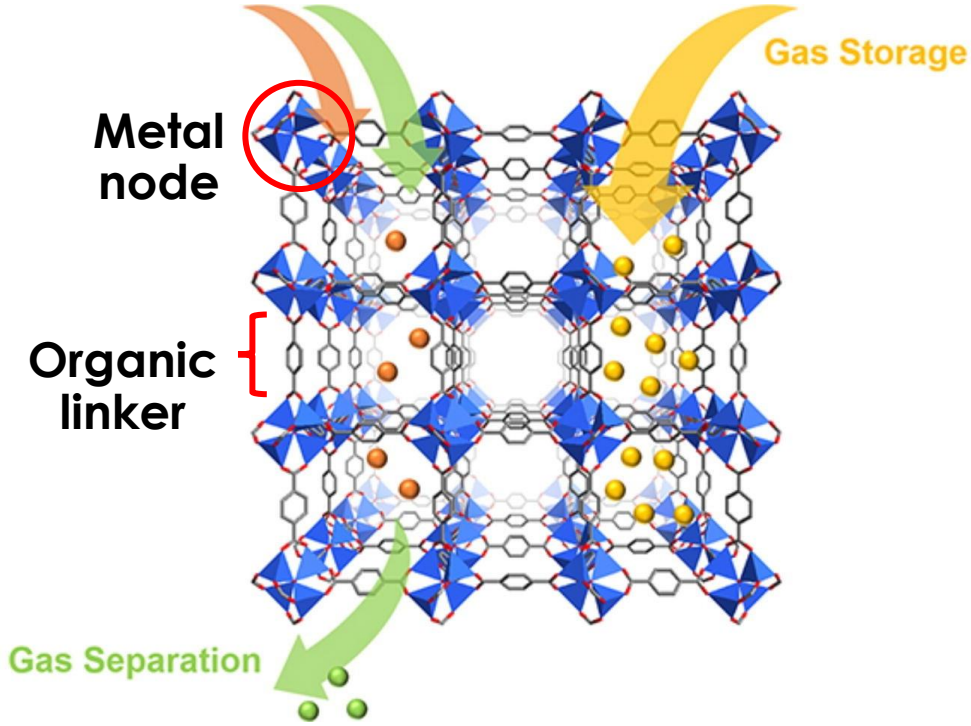
Disclaimer:

This project was funded by the United States Department of Energy, National Energy Technology Laboratory, in part, through a site support contract. Neither the United States Government nor any agency thereof, nor any of their employees, nor the support contractor, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

Liu et al. Science (2021)

Metal Organic Frameworks (MOFs)

- Crystalline, porous materials
- Good for gas storage/separation applications
- Large, diverse class of molecules
  - >100,000 synthesized
  - >500,000 predicted



Li et al. *Materials Today* **2018**, *21* (2), 108–121.

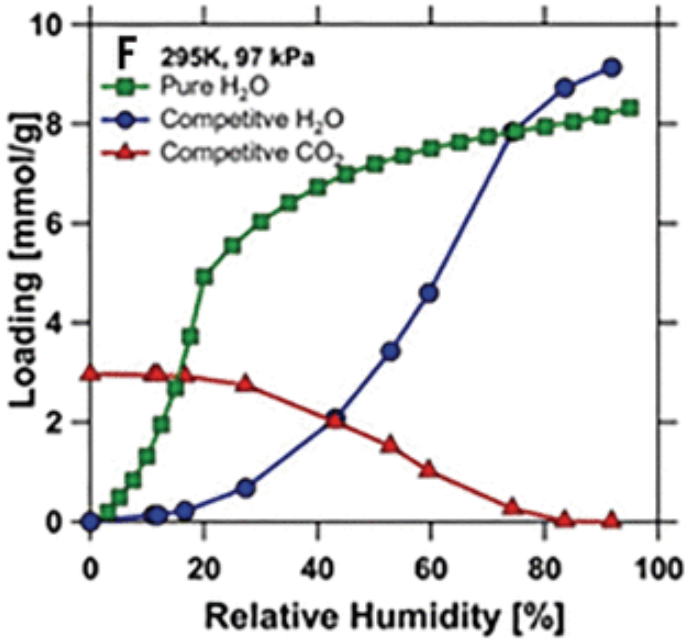
Physisorbent MOFs for CO_2 Capture

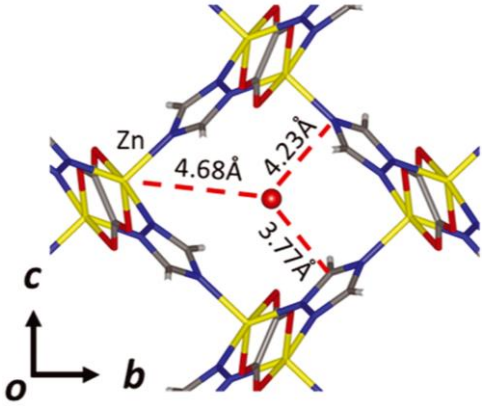
Characteristics of a good capture sorbent

- Strong CO_2 adsorption
- Regenerable
- Not hindered by humidity

Example: CALF-20

- Successfully used for point source
- Little effect of humidity up to 30% RH
- Strong **dispersion** interactions (CO_2 > H_2O)



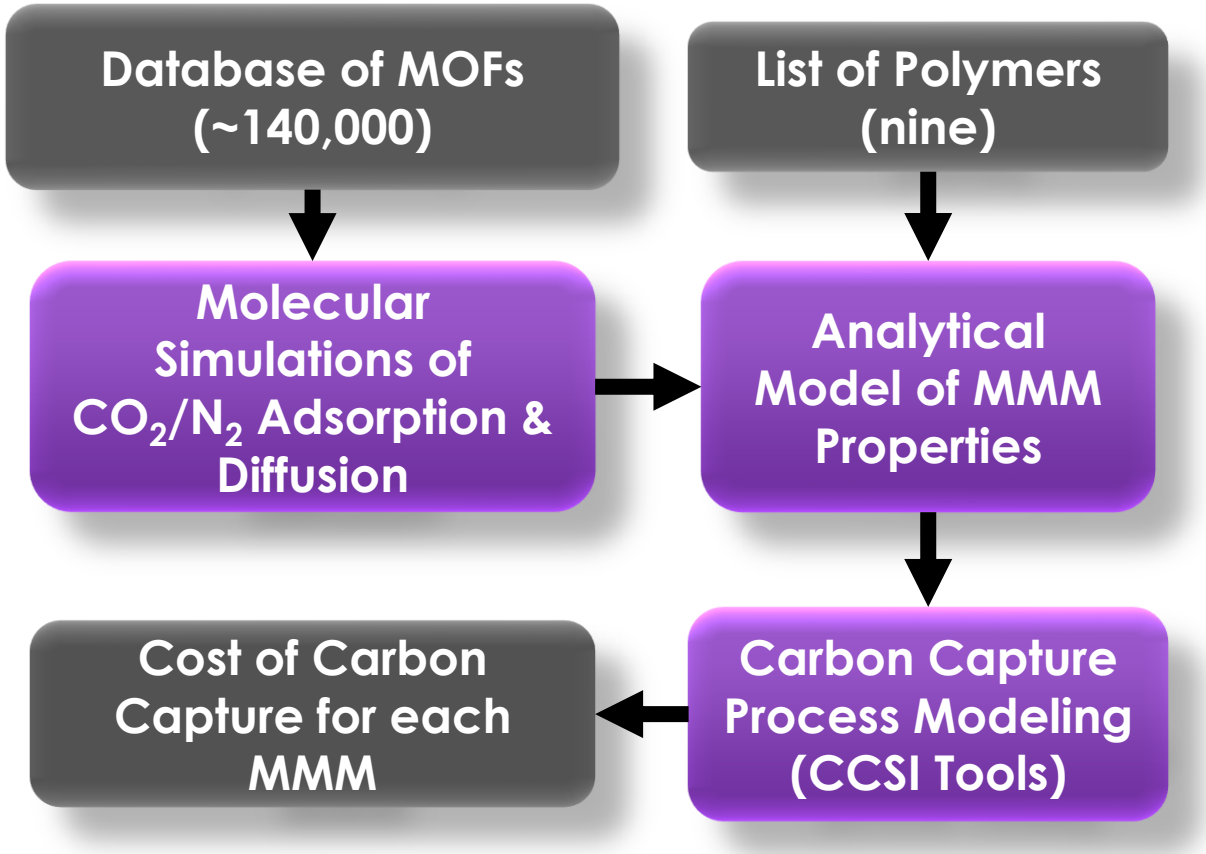


Lin et al., *Science* (2021)

Computational Screening for Mixed Matrix Membranes (MMMs)

NATIONAL ENERGY TECHNOLOGY LABORATORY

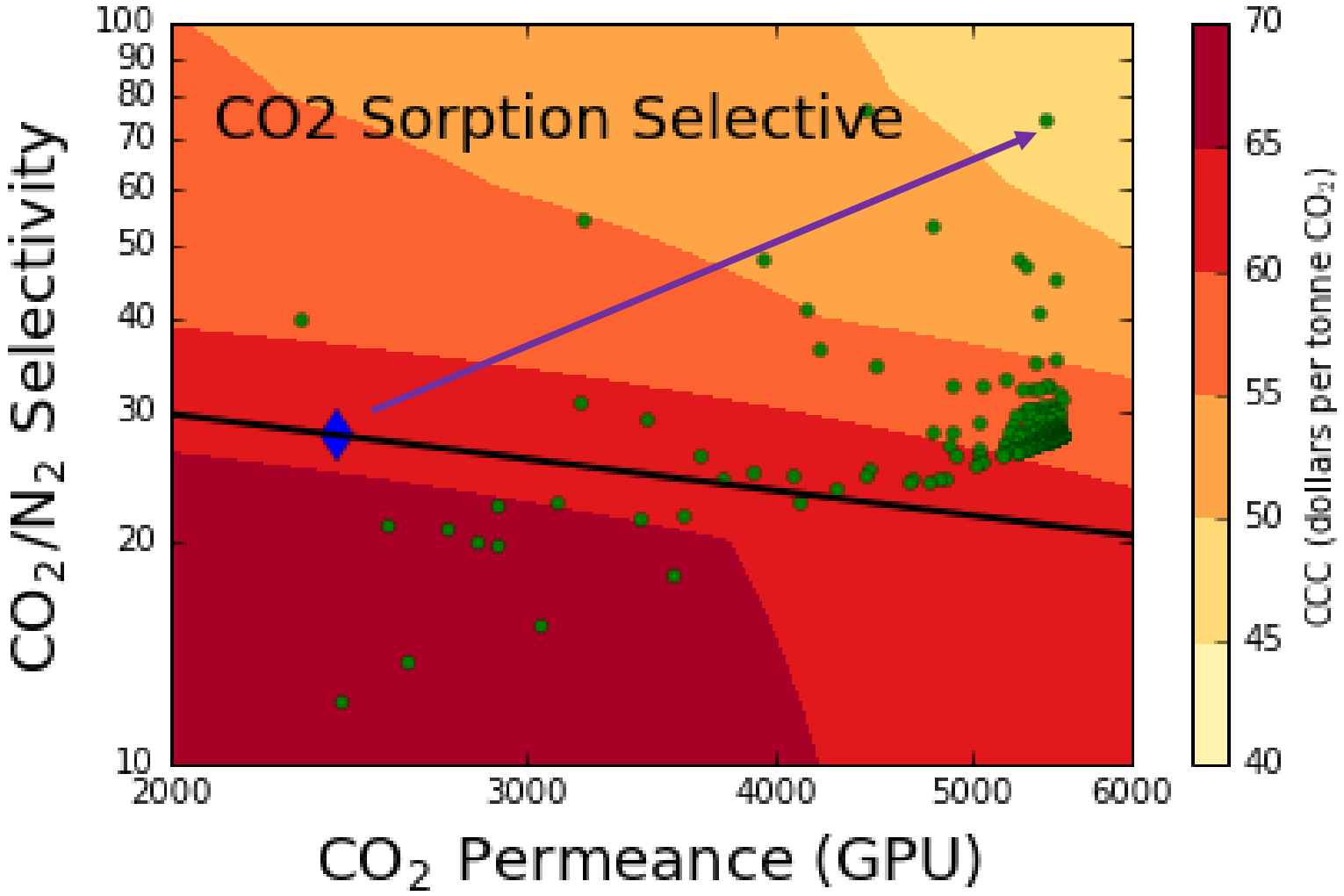
MMMs can be made from Combining MOFs and Polymer Materials



Predicted Cost of Capture Process for >1 million membranes

Samir Budhathoki, Kayode Ajayi, Christopher E. Wilmer

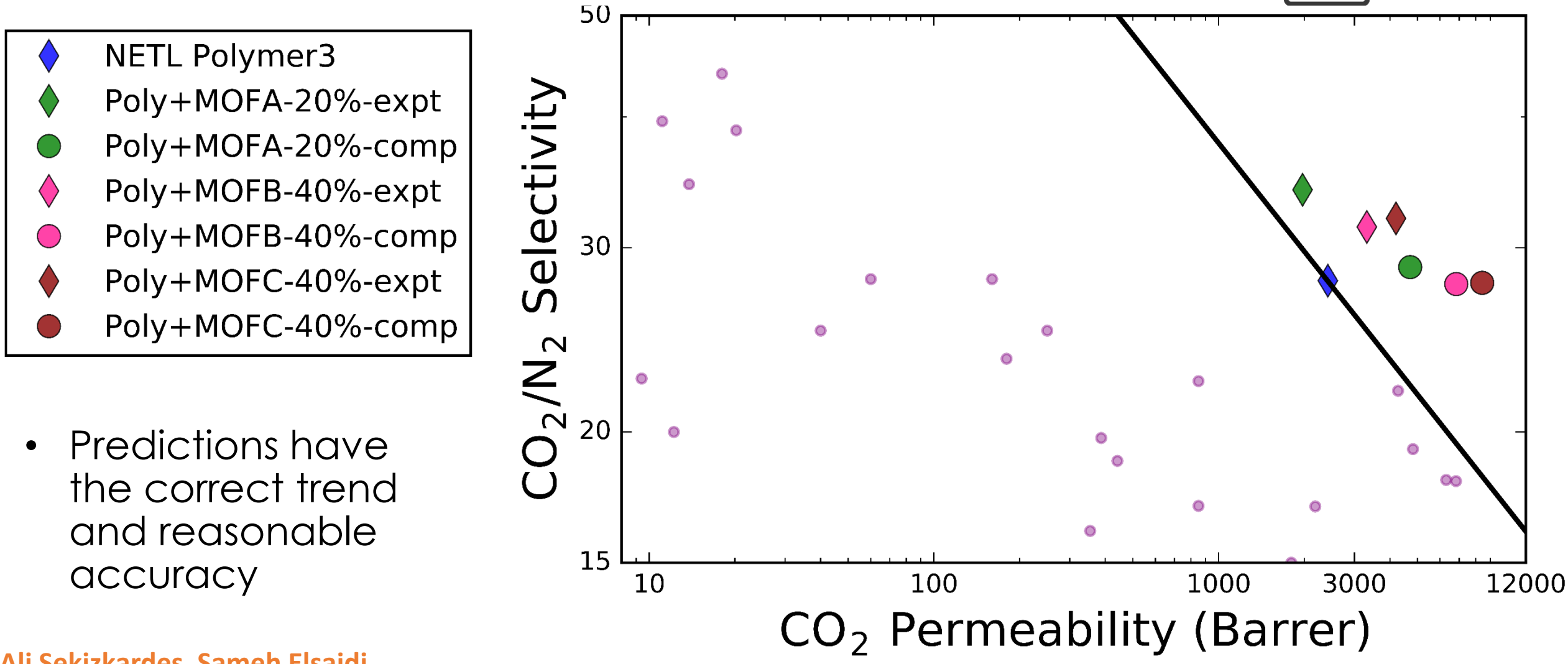
Predictions from High Throughput Computational Screening



- MMMs using NETL Polymer 3 (blend polymer)
- Best MMM in this set:
  - Predicted CCC Reduction from \$62.9 to \$42.7 per tonne CO_2
- Tool for selecting MOFs to pair with polymers

Budhathoki, Ajayi, Steckel, Wilmer, Energy and Environmental Sciences, 2019
Henry's Constants for H_2O in CoRE MOFs courtesy of:Li, S.; Chung, Y. G.; Snurr, R. Q. *Langmuir* **2016**

Comparison with Experiment

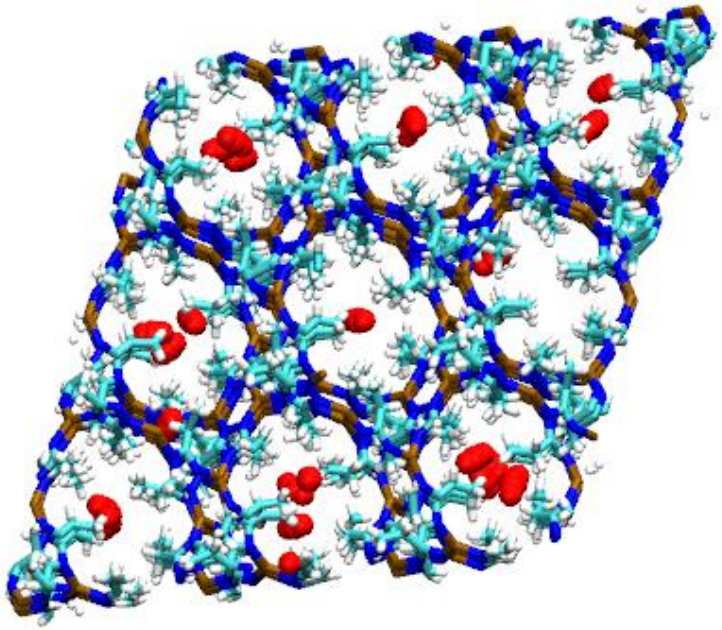


- Predictions have the correct trend and reasonable accuracy

Ali Sekizkardes, Sameh Elsaidi

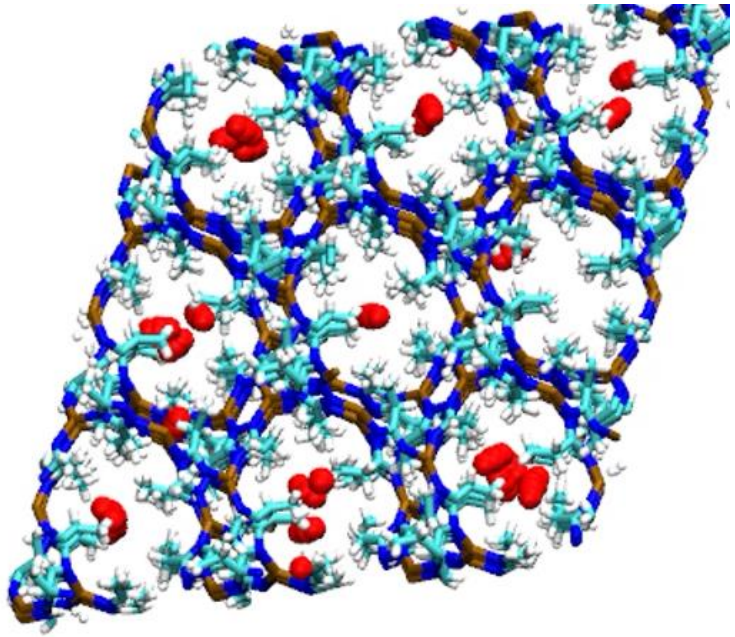
MOF Computational Screening for CO_2 Capture

Past large scale screening studies have modeled MOFs atoms as rigid



Rigid Force Field

- Electrostatic + dispersion
- Low computational cost
- Good for rigid MOFs but most MOFs are flexible
- Easy to obtain

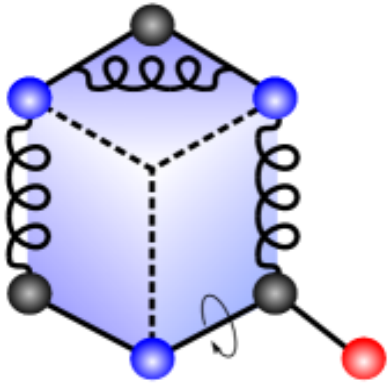


Flexible Force Field

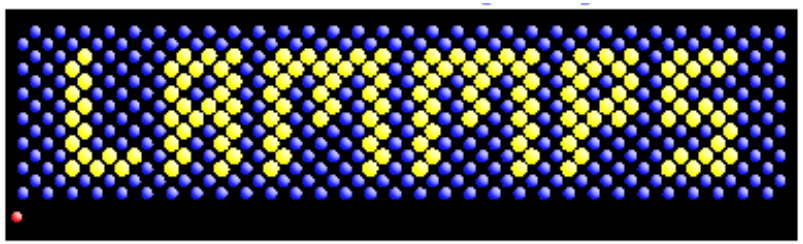
- Electrostatic + dispersion + bond + angle + torsion
- High computational cost
- Accounts for MOF flexibility
- Hard to obtain for a large set of materials

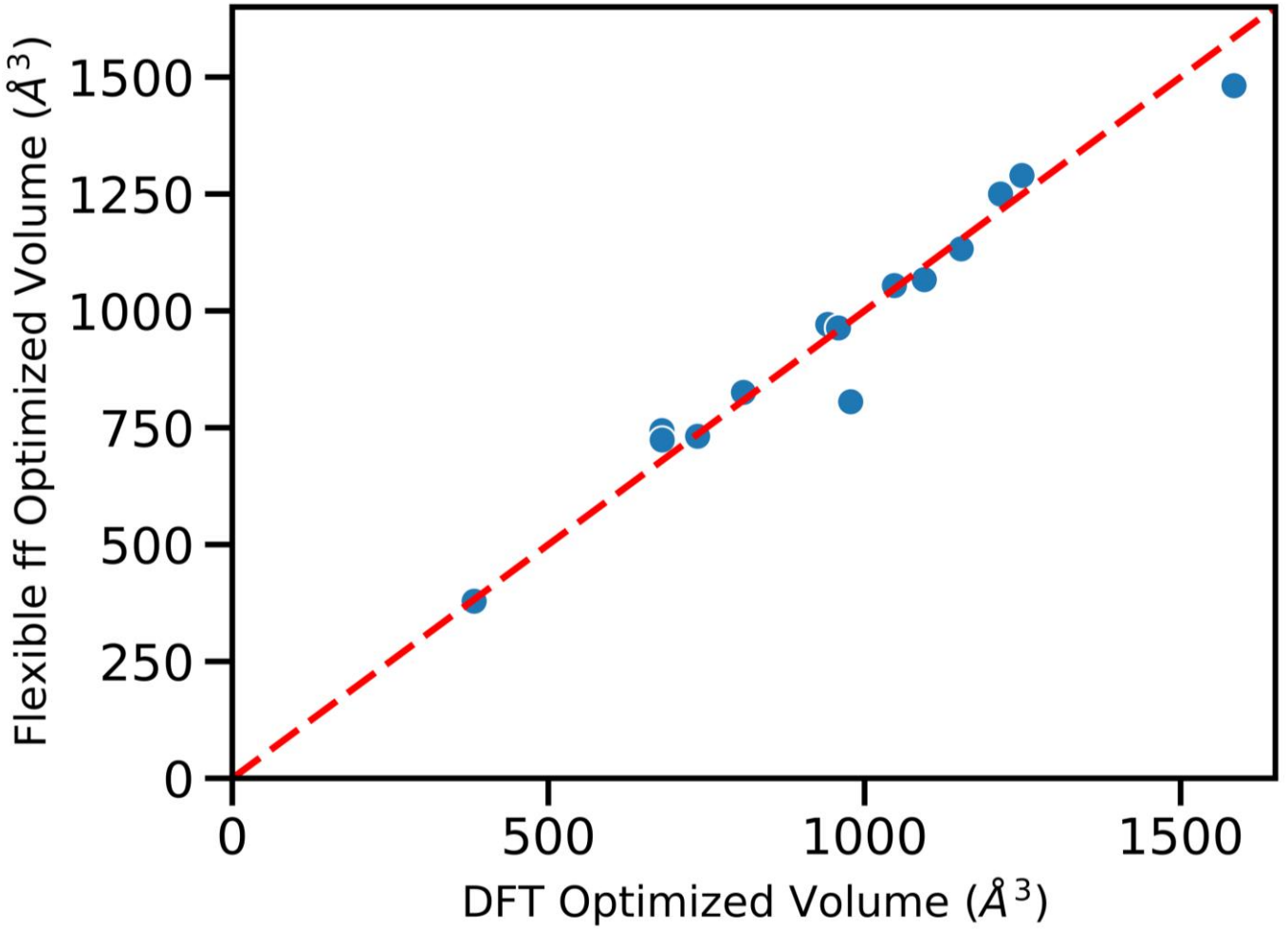
Flexible Force Fields for MOFs

Quick FF used with DFT Vibrational Frequency Calculations to fit flexible FFs



QuickFF

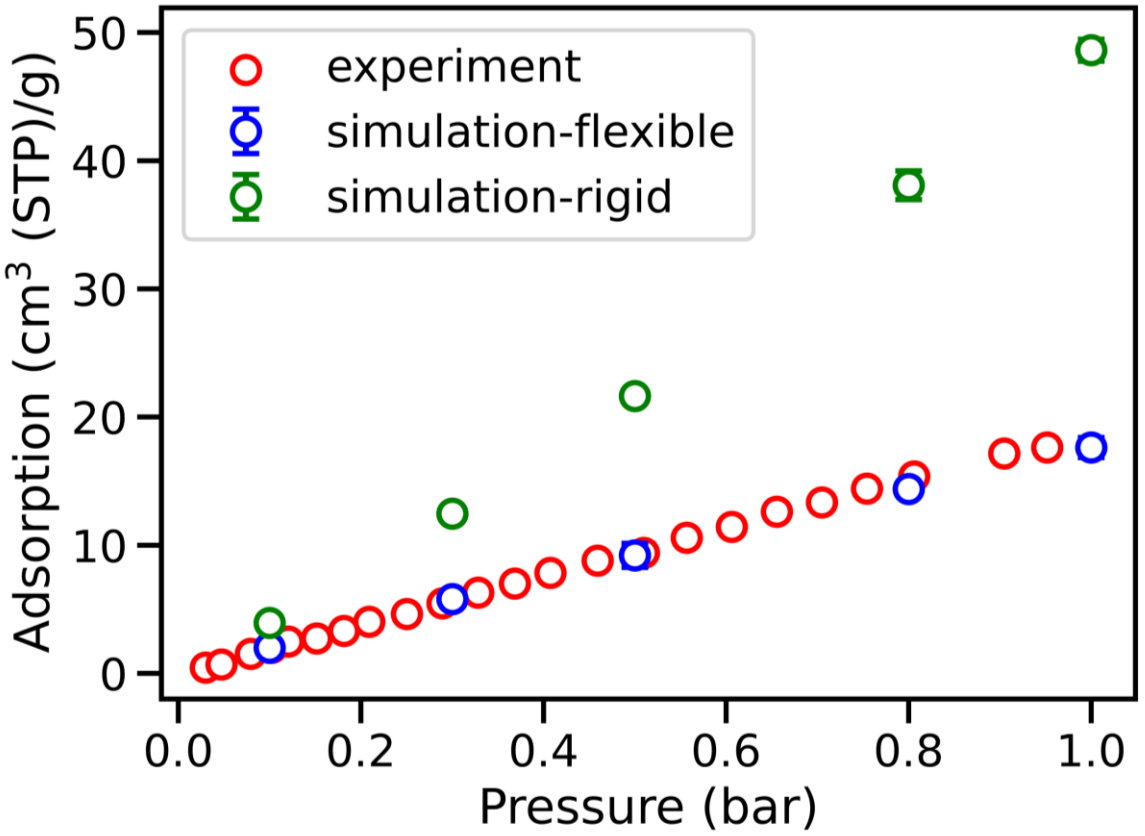


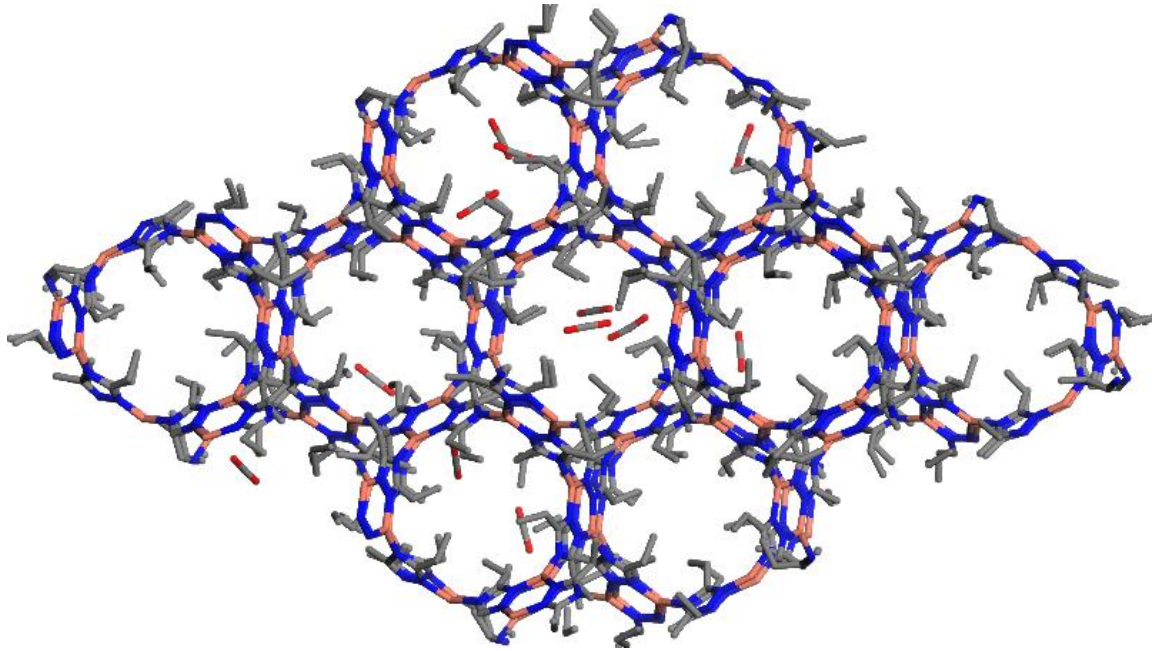


Vanduyfhuys, L. et al. *JCC* **2015**, 36 (13), 1015–1027.
Vanduyfhuys, L. et al. *JCC* **2018**, 39 (16), 999–1011.

Calculated Isotherm: Flexible vs Rigid FF

CO_2 adsorption in MAF-2 (BOGXIF) at 298 K

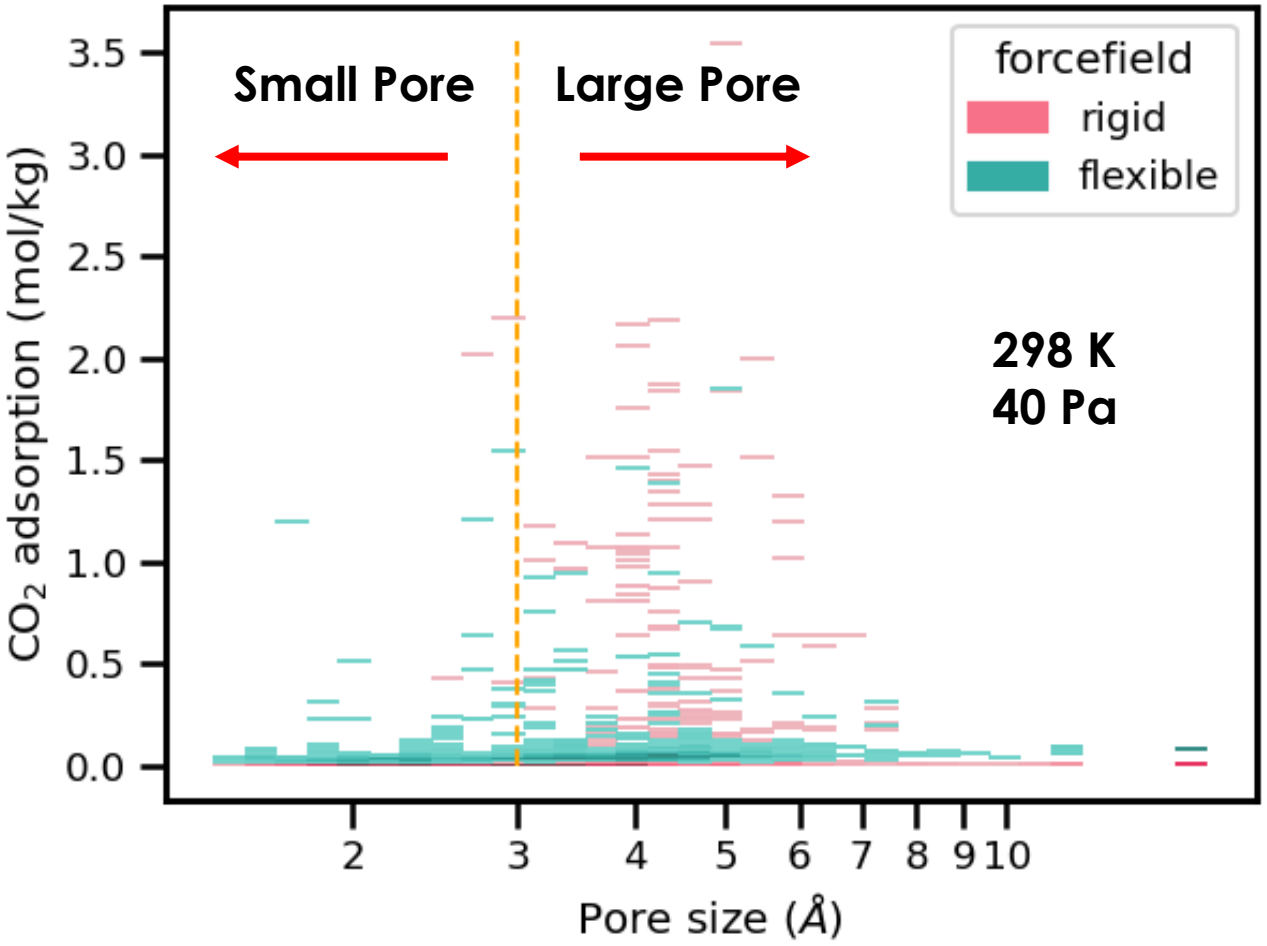




Baucom, T.; Budhathoki, S.; Steckel, J. A. J. Phys. Chem. C **2023**
Zhang, J. P.; Chen, X-M.; *J. Am. Chem. Soc*. **2009**, 131, 15, 5516-5521

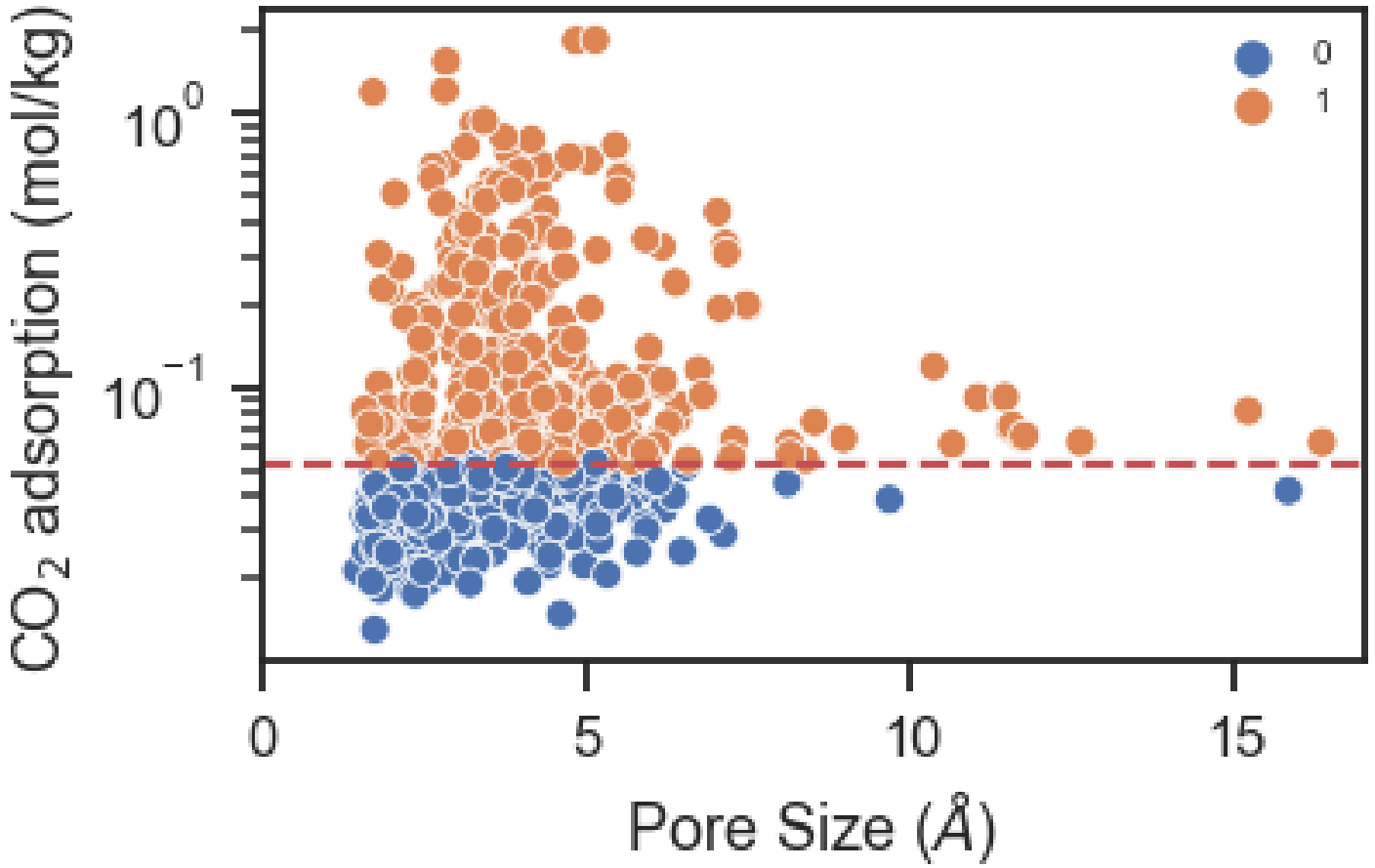
CO_2 Adsorption at Direct Air Capture Conditions

Comparison Between Rigid and Flexible Force Field Results



- Flexible forcefields yield CO_2 adsorption in MOFs with small pore sizes
- Rigid forcefields yield no CO_2 adsorption for MOFs with pore size less than 3 Å in most of the cases
- For some MOFs with pore size greater than 3 Å, rigid forcefields overestimate the CO_2 adsorption

Machine Learning model for MOF classification



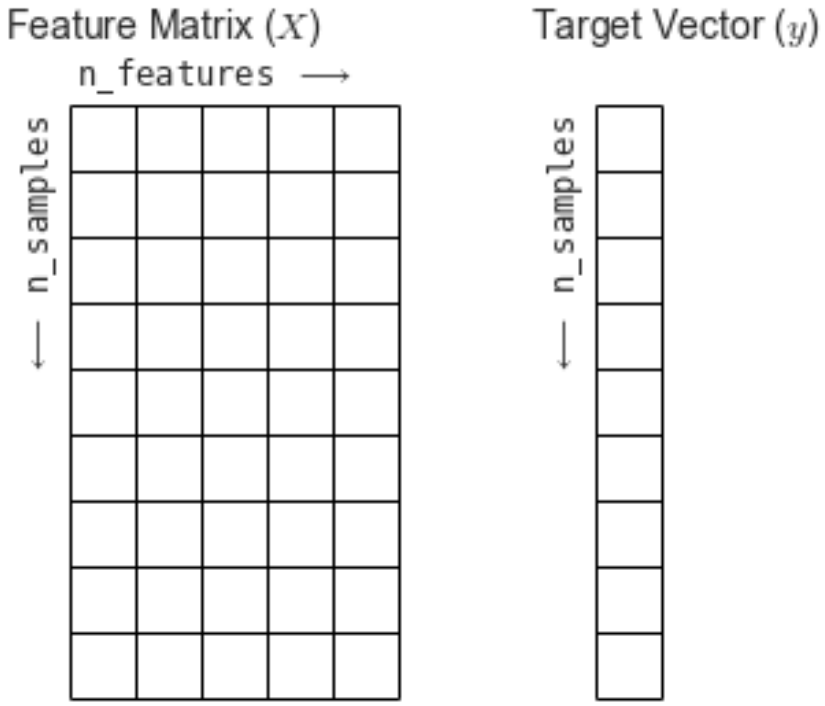
1 – CO_2 ads >= 0.052 mol/kg
0 – CO_2 ads < 0.052 mol/kg

Martin, C.; Budhathoki, S.; Steckel, J. A. Machine Learning Models of CO_2 Sorption and Diffusion in MOFs for Direct Air Capture. *In Preparation* **2023**.

MOF Featurization

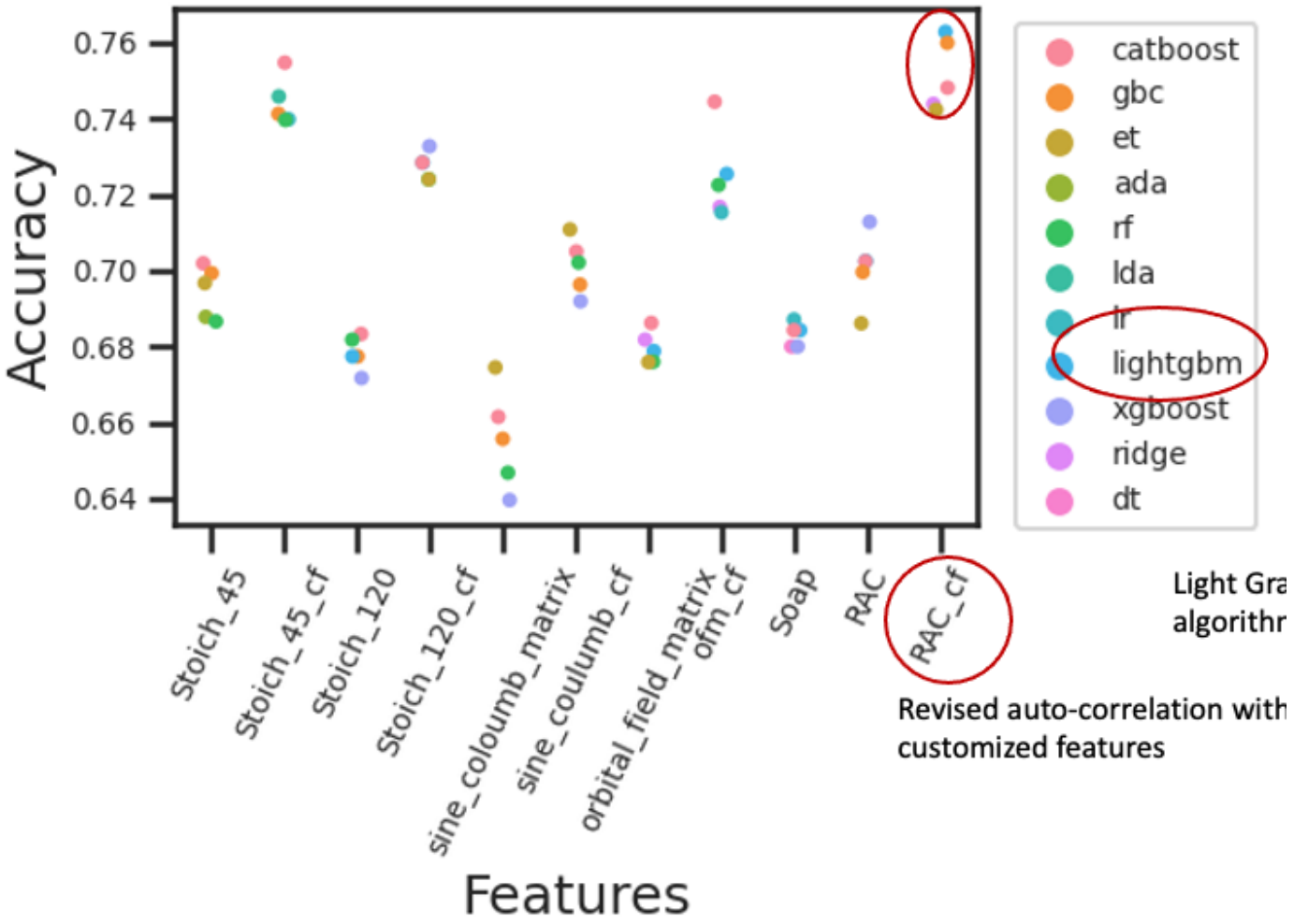
Feature: n-dimensional numerical vector that represents each MOF

1. **Stoichiometric-45**
   - 45 statistical attributes of elemental properties
2. **Stoichiometric-120**
   - 103 attributes describing elemental fractions
   - 7 statistical attributes of elemental properties
3. **Sine Coulomb Matrix**
   - pairwise electrostatic interactions between nuclei
4. **Orbital Field Matrix**
   - distribution of valence electrons
   - interaction of valence subshells between atoms
5. **Smooth Overlap of Atomic Positions (SOAP)**
   - similarity between a pair of local atomic environments
6. **Revised Autocorrelation (RAC) values + Custom features**
   - molecular revised autocorrelation (RAC) values, surface area, volume, density, pore-limiting diameter (PLD), charge difference, epsilon



Matter **2021**, *4* (5), 1578–1597.

Results: ML Classifiers validation results



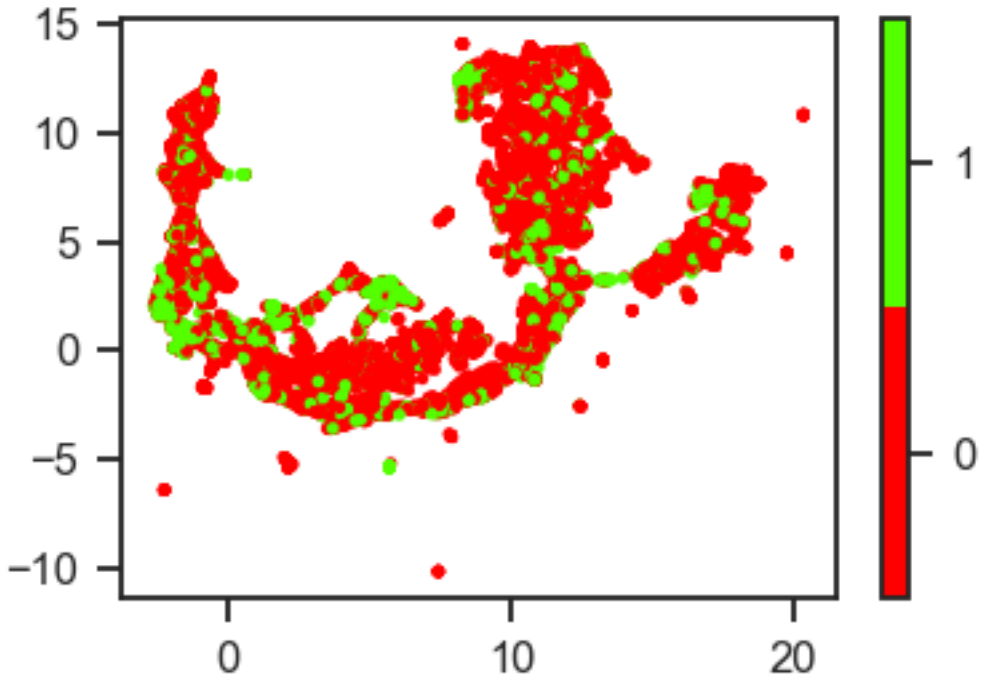
Train size : 80 %
Test size: 20%
Results are averaged from 10 cv folds

- 11 different classification ML models
- All the features were customized (denoted by '_cf' in plot) with geometric features such as surface area, pore size, electrostatics and dispersion term.
- Light Gradient Boosting algorithm performed the best

$$\text{Accuracy} = \frac{True\ Positive + True\ Negative}{True\ Positive + False\ Positive + True\ Negative + False\ Negative} = 76\%$$

Martin, C.; Budhathoki, S.; Steckel, J. A. Machine Learning Models of CO_2 Sorption and Diffusion in MOFs for Direct Air Capture. *In Preparation* **2024**.

Results: Predictions on rest of QMOF Database

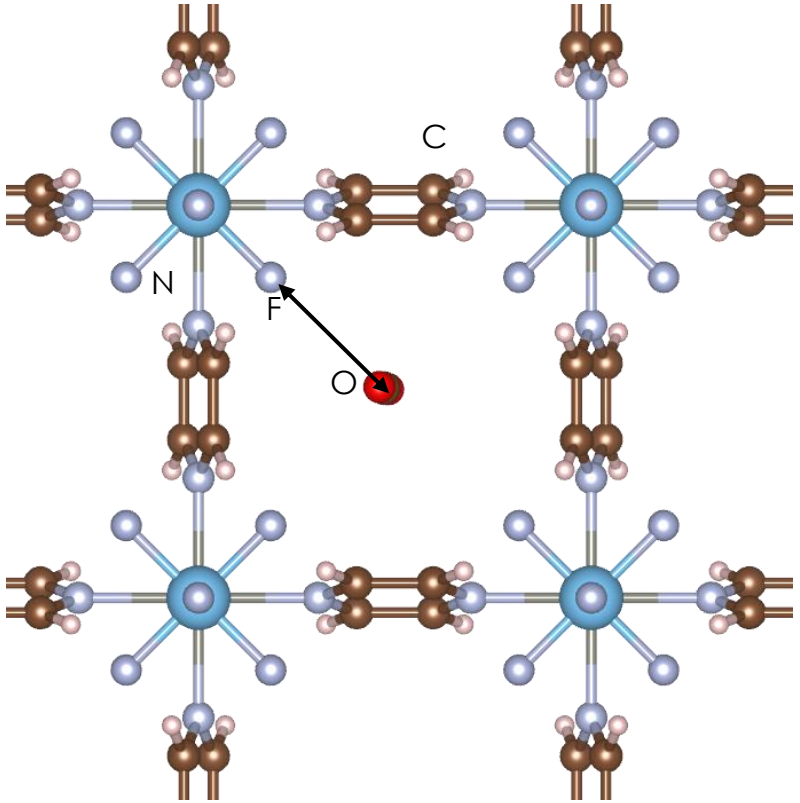


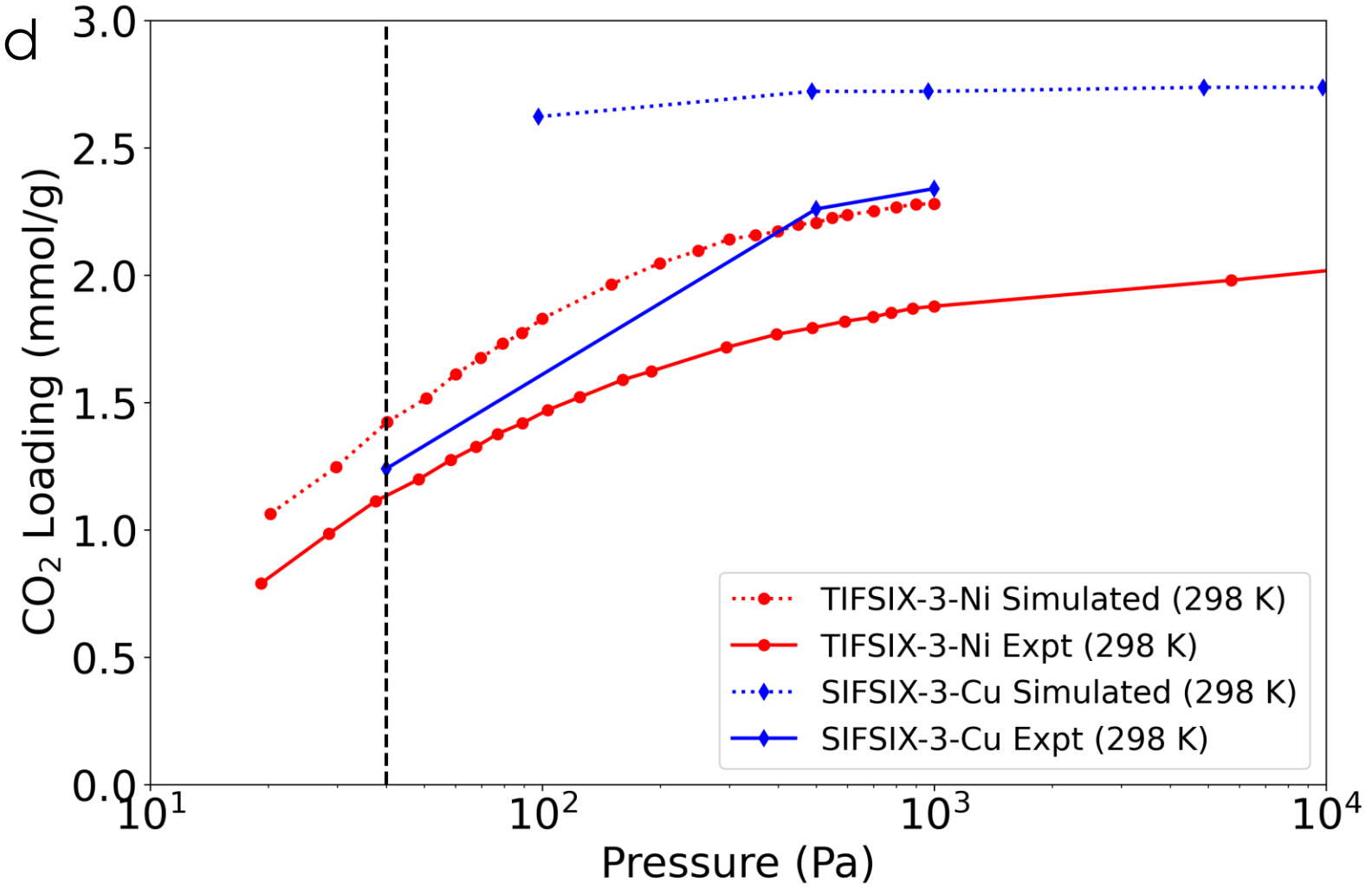
Total MOFs	High adsorbing MOFs (1) (CO_2 adsorption > = 0.05 mol/kg)	Low adsorbing MOFs (0) (CO_2< 0.05 mol/kg)
10,645	2447 (~23 %)	8198 (~77%)

Performance of Classical Force Fields: TIFSIX, SIFSIX

Performance at Low Pressure

CO_2 adsorption overpredicted using rigid force fields



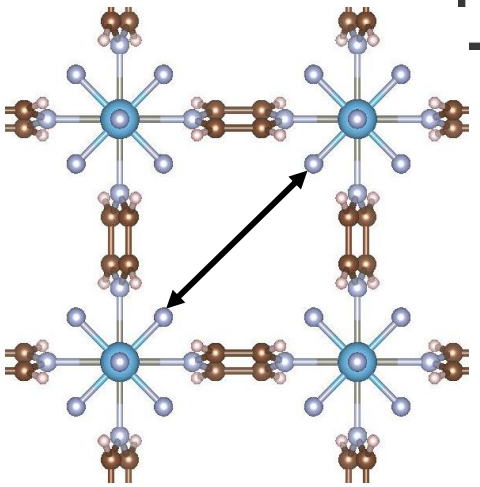


[3]Ullah et al., *Angewandte Chemie* (2022) [4]Low et al. *Energy& Fuels* (2024)
[5]Shekhah et al., *Nature Communications* (2014) [6]Mulcair, *Dissertation* (2017) [7]Forrest et al., *Crystal Growth and Design* (2019)

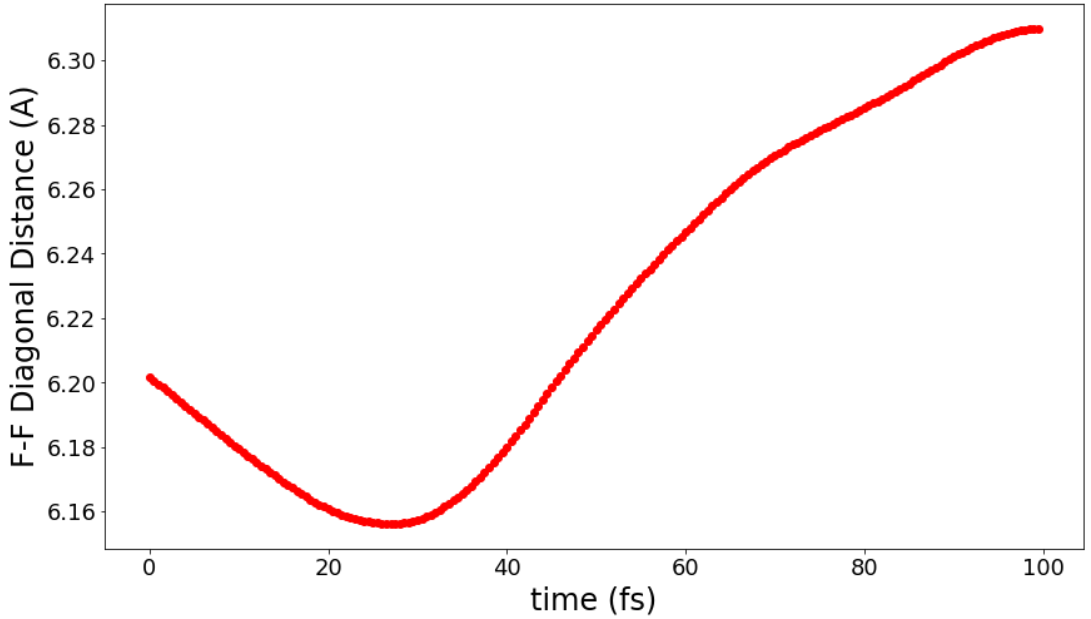
Importance of Flexibility and Accuracy

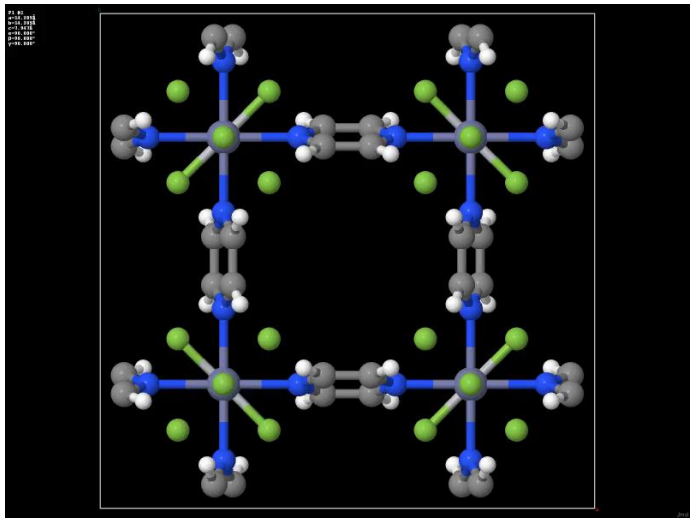
Flexible SiF_6 and TiF_6 can Cause Changes in Window Size

- Poor performance of rigid force fields
  - Overprediction at low P_{CO2}
- DFT molecular dynamics:
  - Significant motion of fluorine atoms
  - Fluorine location affects adsorption strength



F-F diagonal

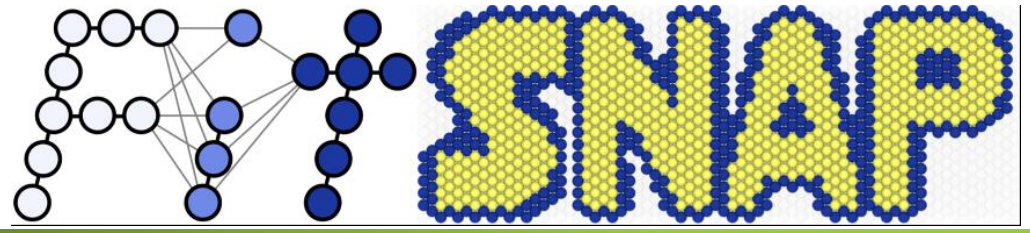




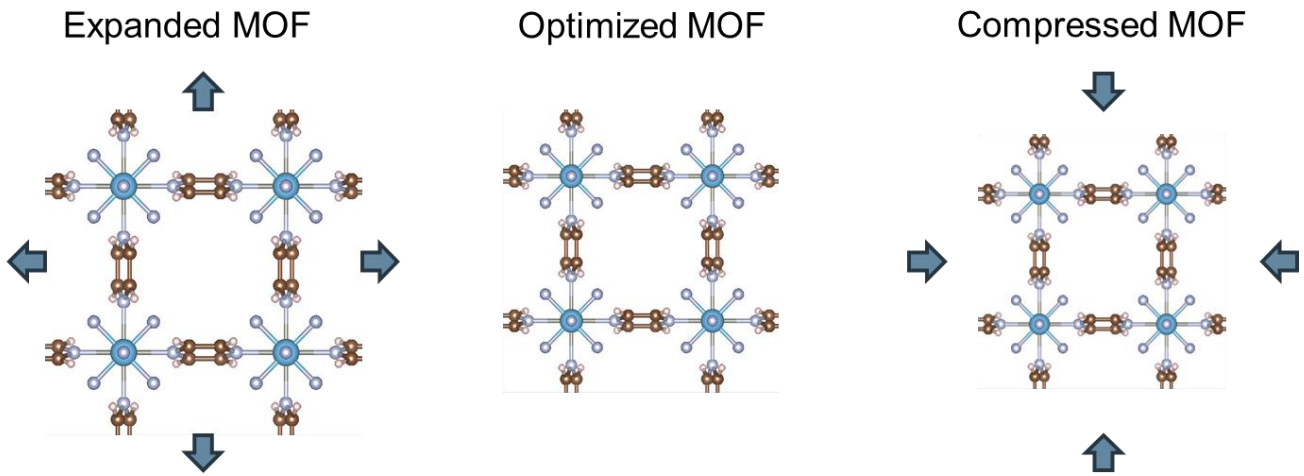
MLFF Training: Spectral Neighbor Analysis Potential

MLFFs Trained on Accurate DFT Data

- **MLFF Method: SNAP**
  - regression model maps local atomic environments to atomic energies and forces
- **Training Configurations**
  - Favorable (low energy) and unfavorable (high energy)
  - MOF with and without CO_2
- **Our Approach**
  - Hybrid SNAP/classical potential
  - MLFF handles short range
  - Classical handles long range



Sample Volume Changes



Sample CO_2 Adsorption

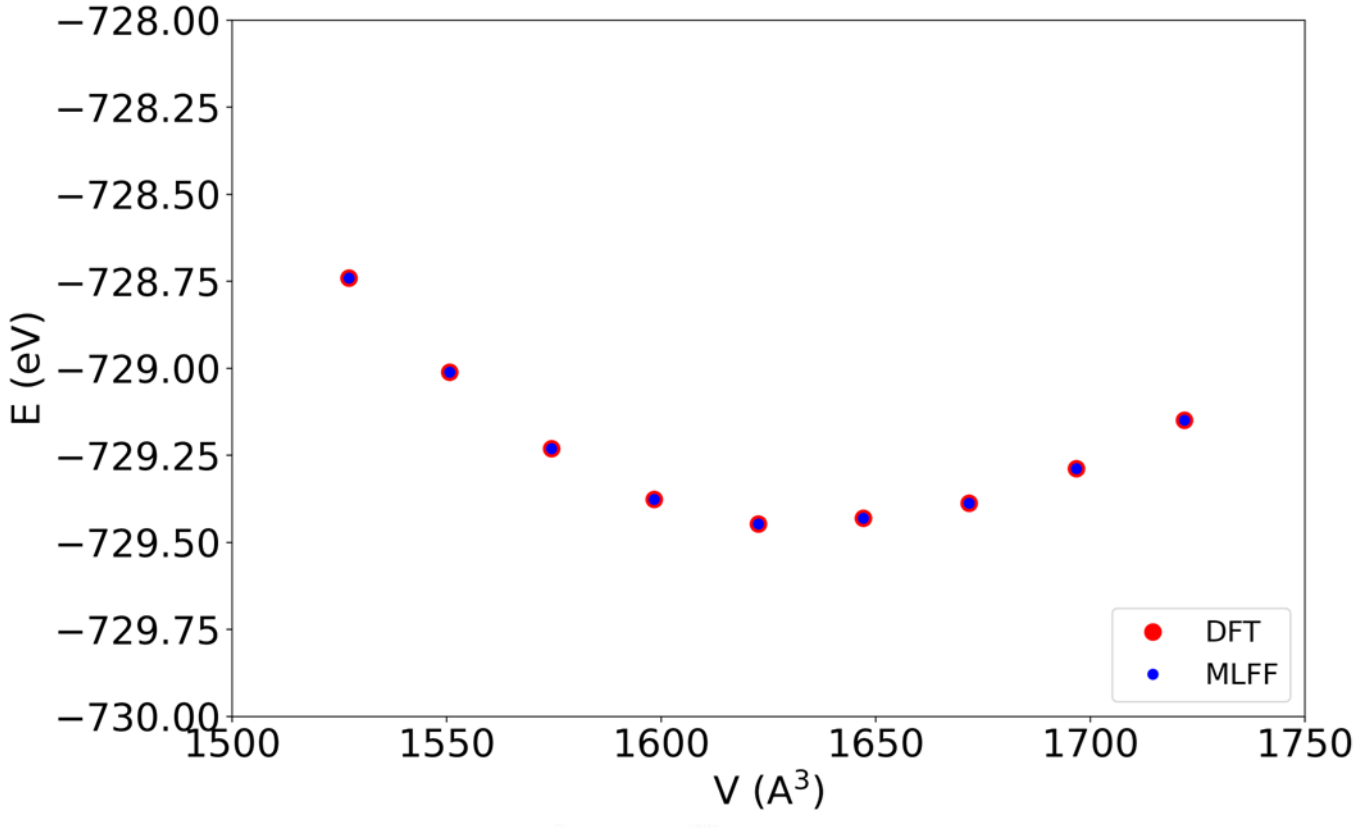
Rohskopf, A. et al. *JOSS* **2023**, *8* (84), 5118.
Thompson, A. P. et al. *JCP* **2015**, *285*, 316–330.

Performance of MLFF Model (Empty)

MLFF Describes MOF Structure and Flexibility (TIFSIX_3_Zn)

- Prediction of energy vs. volume curve is almost perfect (bottom)
  - Related to bulk modulus
- Compared energies/forces in QM-based dynamics calculations, performance is good
  - 7000 training configurations
  - 1800 testing configurations

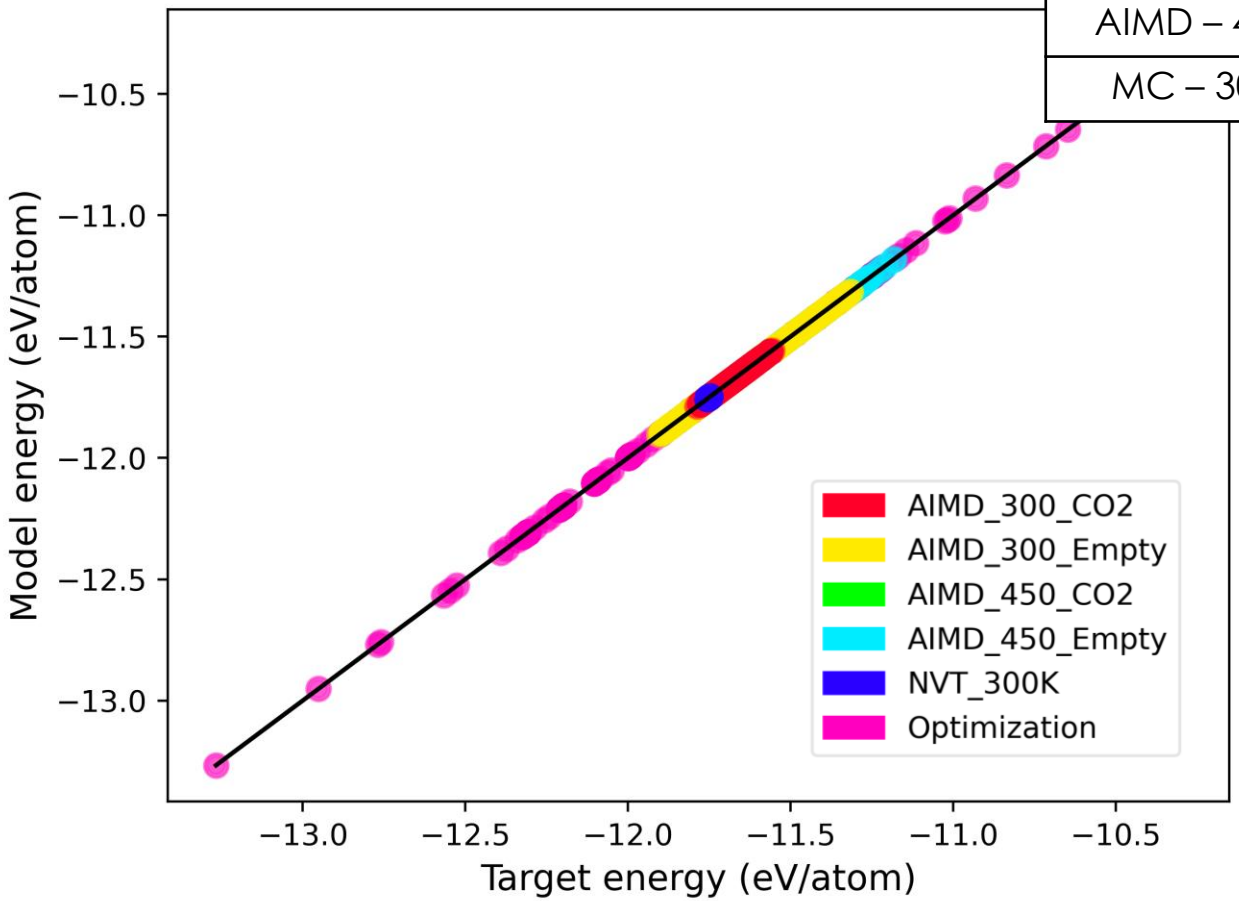
Data Empty MOF (Testing set)	R^2 Energies	R^2 Forces
Structure Optimization	0.998	0.990
AIMD – 300 K	0.991	0.981
AIMD – 450 K	0.990	0.975

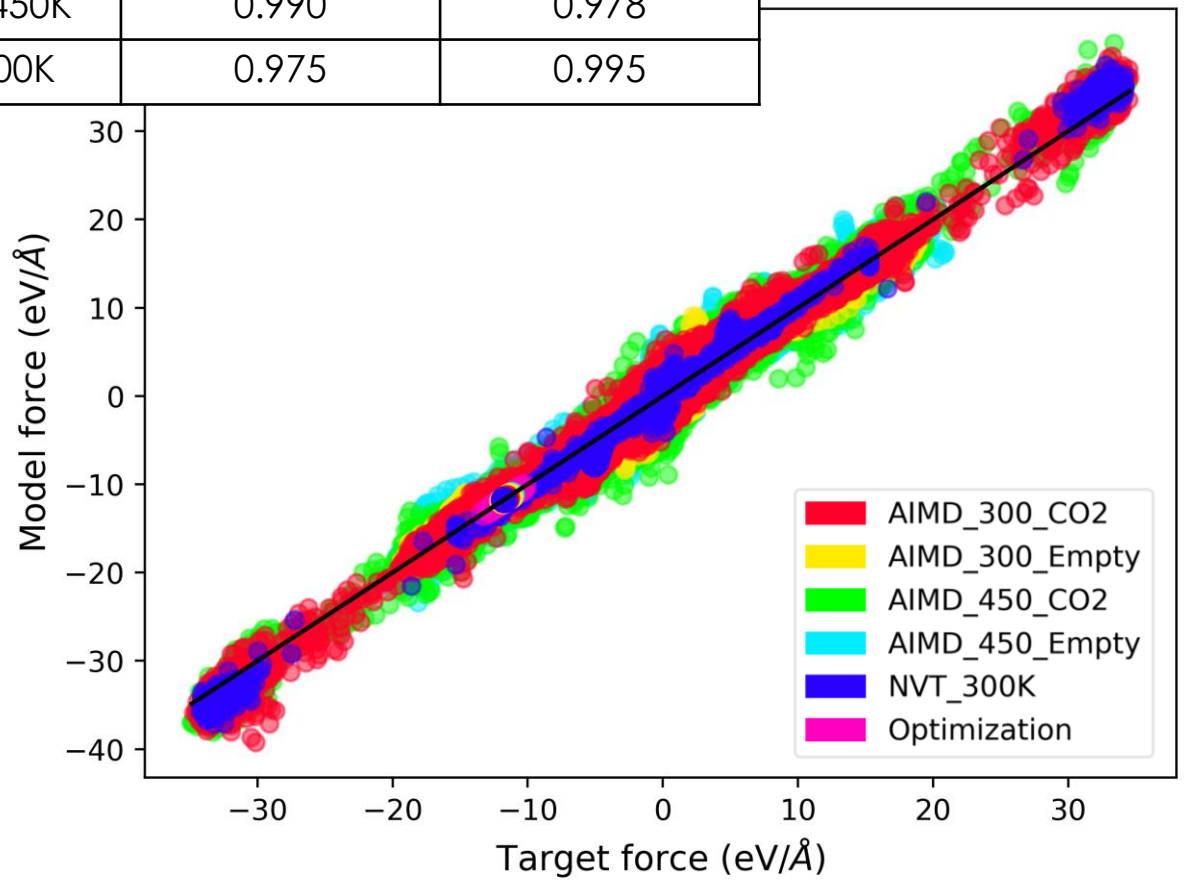


Performance of MLFF: MOF + CO_2

Parity Plots for Energies and Forces

Data MOF + CO_2 (Testing set)	R^2 Energies	R^2 Forces
AIMD – 300K	0.990	0.983
AIMD – 450K	0.990	0.978
MC – 300K	0.975	0.995





Performance of MLFF Force Field: TIFSIX-3-Zn

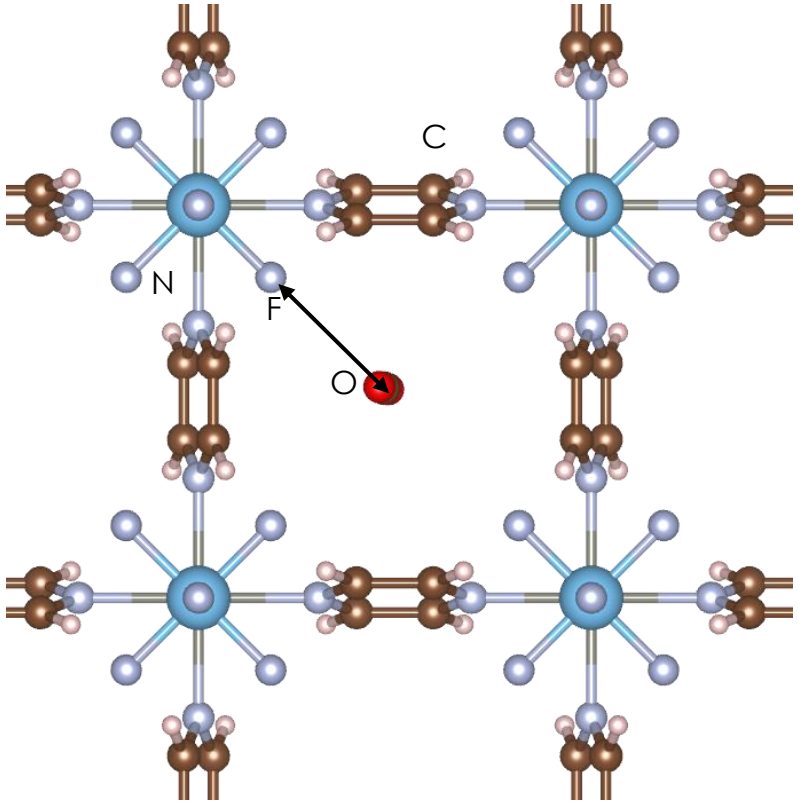
Performance at Low Pressure

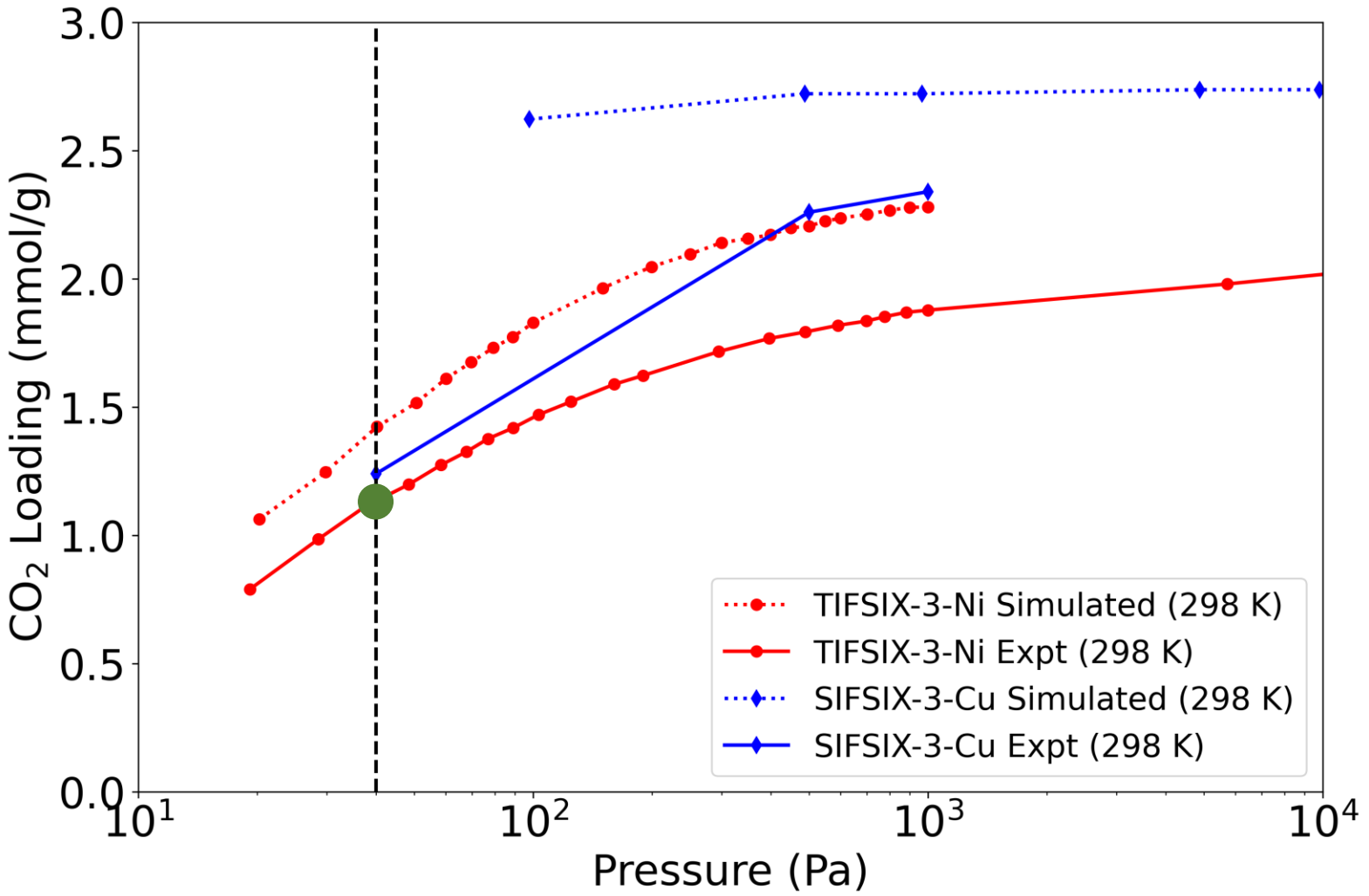
One point on isotherm so far

MLFF Simulation:

TIFSIX-3-Zn adsorbs1.08 mmol/g*

- 400 ppm, 298 K





[3]Ullah et al., *Angewandte Chemie* (2022) [4]Low et al. *Energy& Fuels* (2024)
[5]Shekhah et al., *Nature Communications* (2014) [6]Mulcair, *Dissertation* (2017) [7]Forrest et al., *Crystal Growth and Design* (2019)

Conclusions

- Inclusion of MOFs in polymer membranes can lower the cost of membrane-based CO_2 capture by up to 1/3
- Rigid FF can lead to large overprediction of CO_2 isotherm at low pressure
- MLFF flexible potential accurately describes energies and forces in a MOF loaded with CO_2
- Training on the error of the classical force field improves the MLFF model fit
- TIFSIX-3-Zn $\Delta E_{ads,\ CO2}$, MLFF = -52.9 kJ/mol, DFT -52.1 kJ/mol
- MLFF Prediction: TIFSIX-3-Zn adsorbs 1.08 mmol/g* at 400 ppm, 298 K