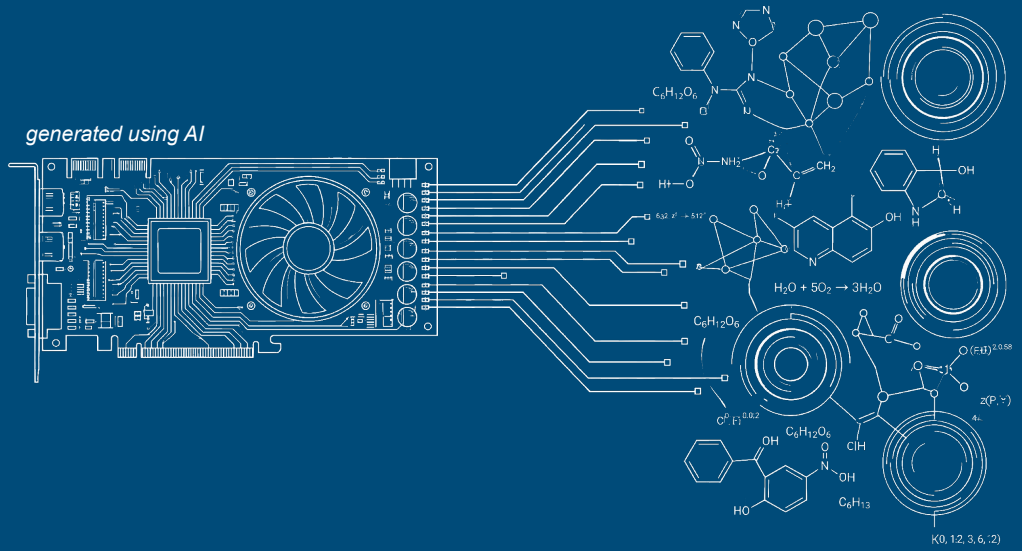
October 28th, 2026



Accelerating Molecular Discovery through GPU Computing

Maxime Ferrer, Thijs Stuyver

PR[AI]RIE AI Fellow

Chimie Paris ParisTech | PSL | GEM | FRANCE 2030 | PR[AI]RIE Paris School

Basis of computational chemistry

Basis of computational chemistry

1 important property
the electronic energy

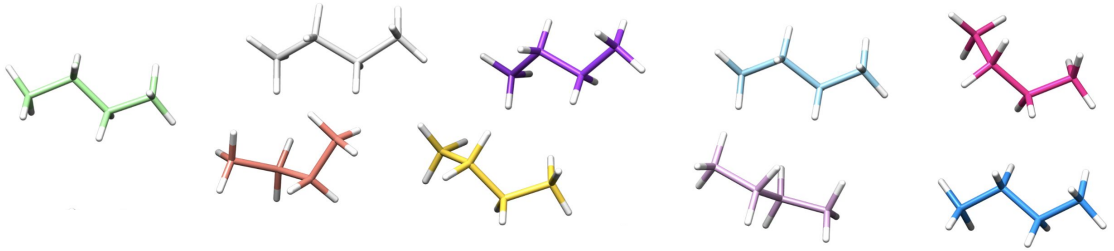
Basis of computational chemistry

1 important property
the electronic energy

electronic energy depends on the atoms in the molecules, the geometry of the molecules => generation of a certain distribution of electrons in molecular orbitals

Basis of computational chemistry

1 important property
the electronic energy

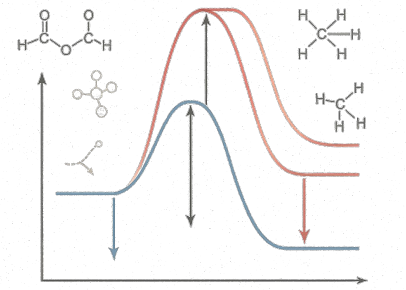


electronic energy depends on the atoms in the molecules, the geometry of the molecules => generation of a certain distribution of electrons in molecular orbitals

- compare the stability of different conformers/tautomers

Basis of computational chemistry

1 important property
the electronic energy



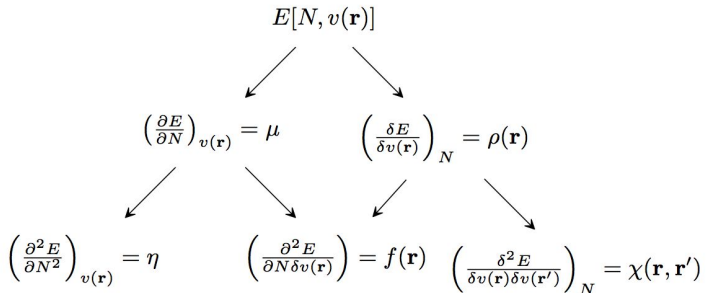
electronic energy depends on the atoms in the molecules, the geometry of the molecules => generation of a certain distribution of electrons in molecular orbitals

- compare the stability of different conformers/tautomers
- compare the reactivity of different systems by computing their reaction energy profiles

Basis of computational chemistry

1 important property
the electronic energy

electronic energy depends on the atoms in the molecules, the geometry of the molecules => generation of a certain distribution of electrons in molecular orbitals



- compare the stability of different conformers/tautomers
- compare the reactivity of different systems by computing their reaction energy profiles
- extract from the converged structure chemical properties (acidity, basicity, electronegativity …)

Basis of computational chemistry

1 important property
the electronic energy

→ How to compute it ?

Basis of computational chemistry

1 important property
the electronic energy

→ How to compute it ?

Traditionally:

"New" approaches:

Basis of computational chemistry

1 important property
the electronic energy → How to compute it ?

Traditionally:

- Wavefunction based theory: HF, MP2, CCSD(T)

$$H\Psi=E\Psi$$

"New" approaches:

Basis of computational chemistry

1 important property
the electronic energy → How to compute it ?

Traditionally:

- Wavefunction based theory: HF, MP2, CCSD(T)

$$H\Psi=E\Psi$$

- Density based theory: DFT

$$E[\rho]=F[\rho]+\int V(r)\rho(r)d^3r$$

"New" approaches:

Basis of computational chemistry

1 important property
the electronic energy → How to compute it ?

Traditionally:

- Wavefunction based theory: HF, MP2, CCSD(T)

$$H\Psi=E\Psi$$

- Density based theory: DFT

$$E[\rho]=F[\rho]+\int V(r)\rho(r)d^3r$$

\+ most accurate description

\- computationally expensive

"New" approaches:

Basis of computational chemistry

1 important property
the electronic energy

→ How to compute it ?

Traditionally:

- Wavefunction based theory: HF, MP2, CCSD(T)

$$H\Psi=E\Psi$$

- Density based theory: DFT

$$E[\rho]=F[\rho]+\int V(r)\rho(r)d^3r$$

\+ most accurate description

\- computationally expensive

"New" approaches:

- Semi-empirical methods + Force Field: GNF2-xTB, gxtb, AMBER FF, GAFF

  E=sum of classical terms or usage of a parameterized Hamiltonian

Basis of computational chemistry

1 important property
the electronic energy

→ How to compute it ?

Traditionally:

- Wavefunction based theory: HF, MP2, CCSD(T)

$$H\Psi=E\Psi$$

- Density based theory: DFT

$$E[\rho]=F[\rho]+\int V(r)\rho(r)d^3r$$

\+ most accurate description

\- computationally expensive

"New" approaches:

- Semi-empirical methods + Force Field: GNF2-xTB, gxtb, AMBER FF, GAFF

  E=sum of classical terms or usage of a parameterized Hamiltonian

\+ fast, a fraction of DFT

\- lack of accuracy

Basis of computational chemistry

1 important property
the electronic energy

→ How to compute it ?

Traditionally:

- Wavefunction based theory: HF, MP2, CCSD(T)

$$H\Psi=E\Psi$$

- Density based theory: DFT

$$E[\rho]=F[\rho]+\int V(r)\rho(r)d^3r$$

\+ most accurate description

\- computationally expensive

"New" approaches:

- Semi-empirical methods + Force Field: GNF2-xTB, gxtb, AMBER FF, GAFF

  E=sum of classical terms or usage of a parameterized Hamiltonian

\+ fast, a fraction of DFT

\- lack of accuracy

- Machine Learning Interatomic Potentials (MLIP)

Machine Learning Interatomic Potential (MLIP)

1 important property
the electronic energy

→ How to compute it ?

MLIP

Machine Learning Interatomic Potential (MLIP)

1 important property
the electronic energy

→ How to compute it ?

MLIP

UMA: A Family of Universal Models for Atoms

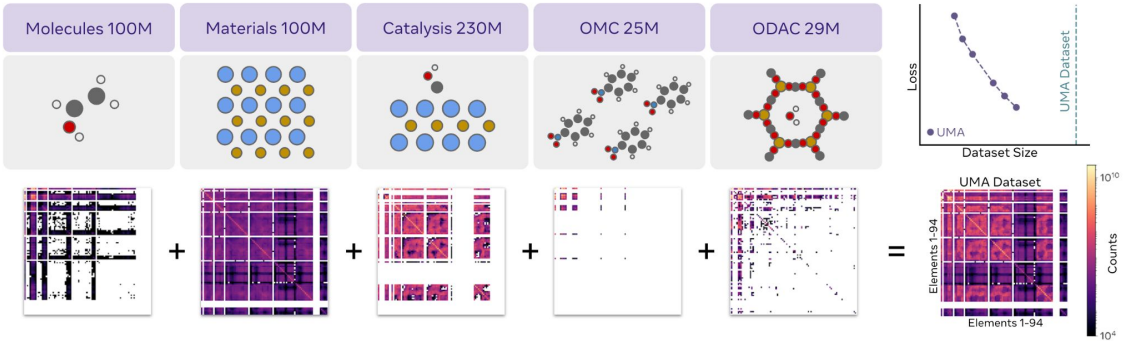
Brandon M. Wood[1,†,*], **Misko Dzamba**[1,*], **Xiang Fu**[1,*], **Meng Gao**[1,*], **Muhammed Shuaibi**[1,*], **Luis Barroso-Luque**[1], **Kareem Abdelmaqsoud**[2], **Vahe Gharakhanyan**[1], **John R. Kitchin**[2], **Daniel S. Levine**[1], **Kyle Michel**[1], **Anuroop Sriram**[1], **Taco Cohen**[1], **Abhishek Das**[1], **Ammar Rizvi**[1], **Sushree Jagriti Sahoo**[1], **Zachary W. Ulissi**[1], **C. Lawrence Zitnick**[1,†]

[1]FAIR at Meta, [2]Department of Chemical Engineering, Carnegie Mellon University
[*]Co-first Author, [†]Co-corresponding Author

The ability to quickly and accurately compute properties from atomic simulations is critical for advancing a large number of applications in chemistry and materials science including drug discovery, energy storage, and semiconductor manufacturing. To address this need, Meta FAIR presents a family of Universal Models for Atoms (UMA), designed to push the frontier of speed, accuracy, and generalization. UMA models are trained on half a billion unique 3D atomic structures (the largest training runs to date) by compiling data across multiple chemical domains, e.g. molecules, materials, and catalysts. We develop empirical scaling laws to help understand how to increase model capacity alongside dataset size to achieve the best accuracy. The UMA small and medium models utilize a novel architectural design we refer to as mixture of linear experts that enables increasing model capacity without sacrificing speed. For example, UMA-medium has 1.4B parameters but only ~50M active parameters per atomic structure. We evaluate UMA models on a diverse set of applications across multiple domains and find that, remarkably, a single model without any fine-tuning can perform similarly or better than specialized models. We are releasing the UMA code, weights, and associated data to accelerate computational workflows and enable the community to continue to build increasingly capable AI models.

Models: https://huggingface.co/facebook/UMA
Code: https://github.com/facebookresearch/fairchem
Correspondence: B.M.W. (bmwood@meta.com), C.L.Z. (zitnick@meta.com)

Meta



from https://huggingface.co/facebook/UMA

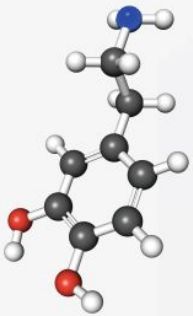
Machine Learning Interatomic Potential (MLIP) - UMA

1 important property
the electronic energy

→ How to compute it ?

MLIP

UMA-OMol illustration



modified from https://www.geeksforgeeks.org/deep-learning/what-are-graph-neural-networks/

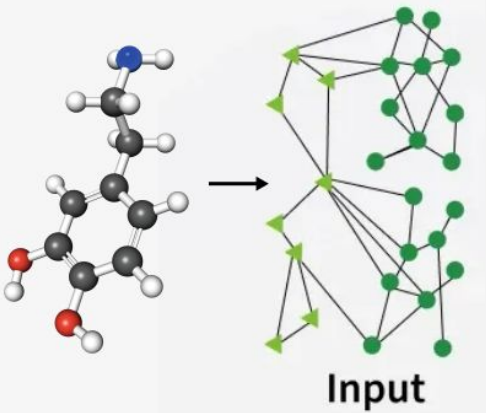
Machine Learning Interatomic Potential (MLIP) - UMA

1 important property
the electronic energy

→ How to compute it ?

MLIP

UMA-OMol illustration



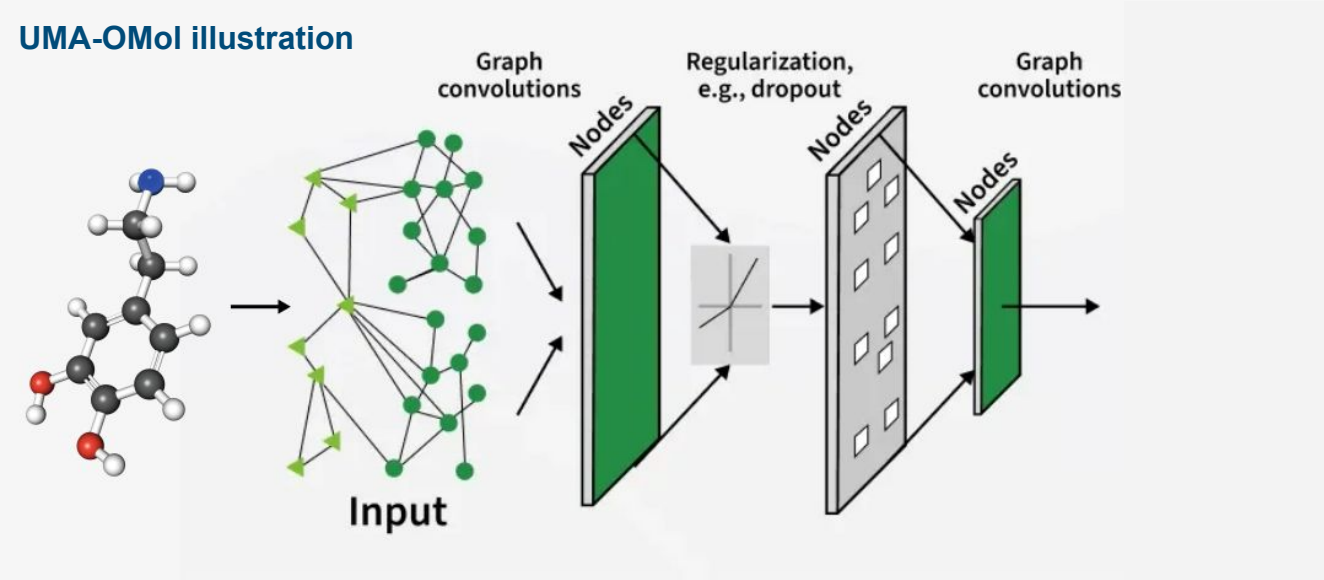
modified from https://www.geeksforgeeks.org/deep-learning/what-are-graph-neural-networks/

Machine Learning Interatomic Potential (MLIP) - UMA

1 important property
the electronic energy

→ How to compute it ?

MLIP



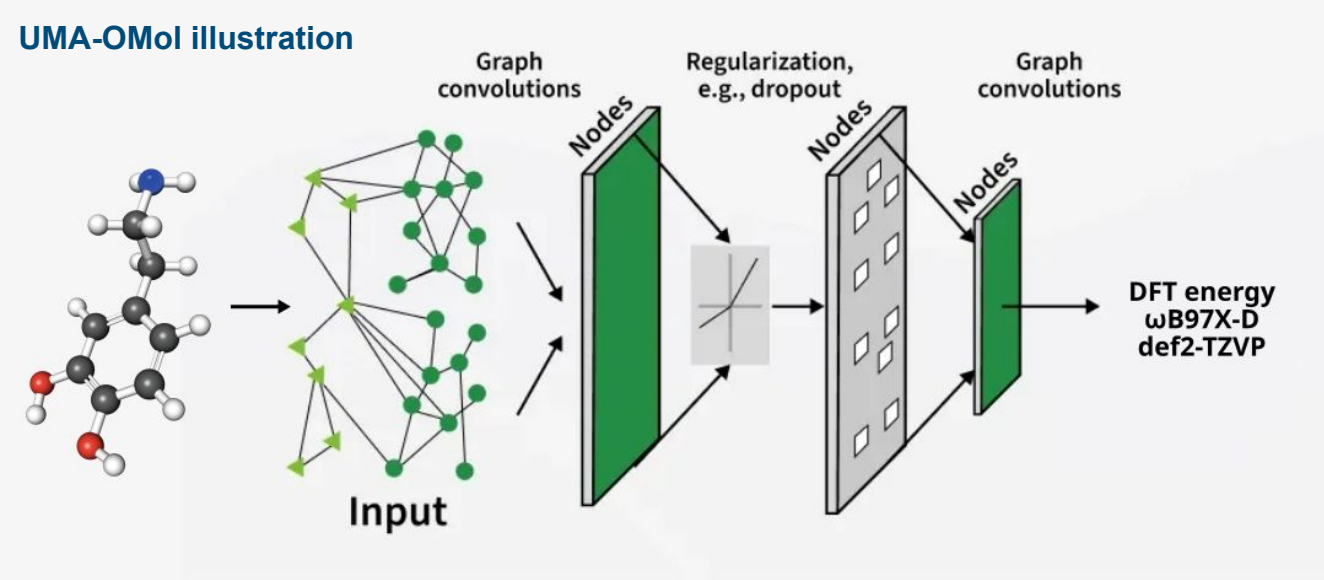
modified from https://www.geeksforgeeks.org/deep-learning/what-are-graph-neural-networks/

Machine Learning Interatomic Potential (MLIP) - UMA

1 important property
the electronic energy

→ How to compute it ?

MLIP



modified from https://www.geeksforgeeks.org/deep-learning/what-are-graph-neural-networks/

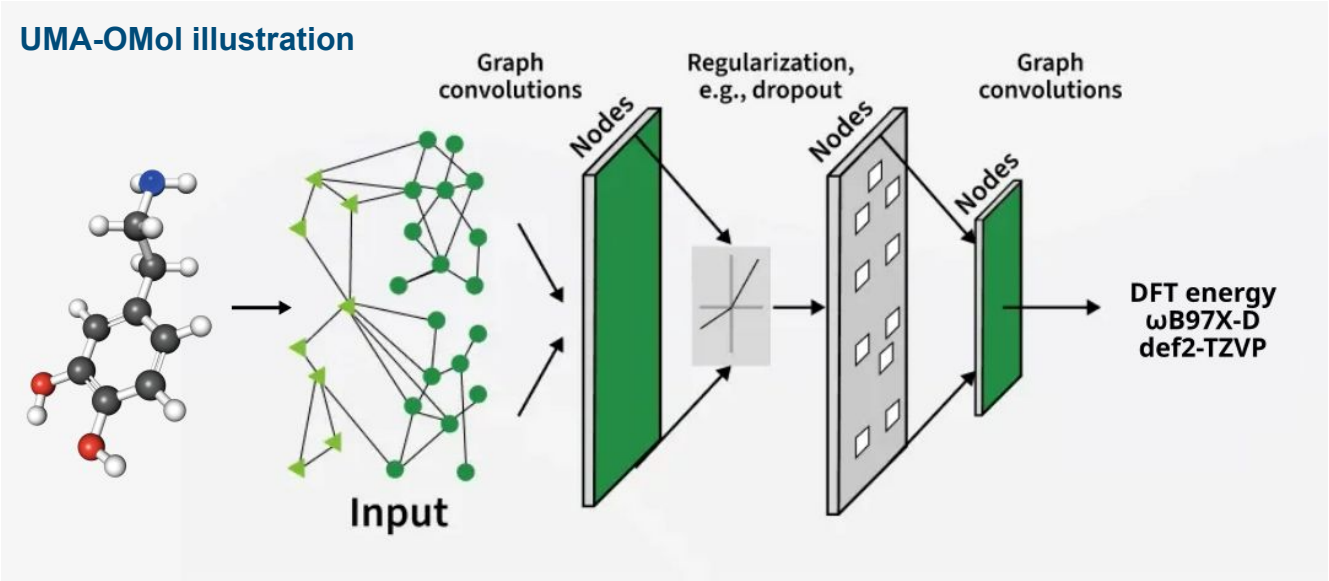
Machine Learning Interatomic Potential (MLIP) - UMA

1 important property
the electronic energy

→ How to compute it ?

MLIP

UMA-OMol illustration



Quick benchmark in the case of FLPs (200 systems, TS and minima) vs DFT:

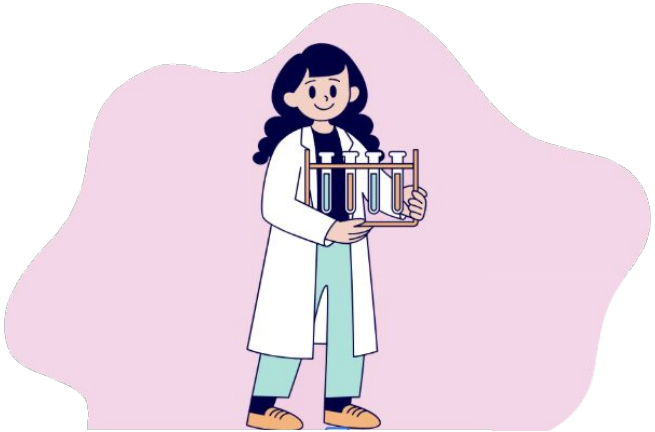
$R^2 > 0.95$
MAE < 2 kcal/mol
RMSE ~ 3 kcal/mol

modified from https://www.geeksforgeeks.org/deep-learning/what-are-graph-neural-networks/

Exploration of Chemical Spaces



Exploration of Chemical Spaces



PREVIOUS KNOWLEDGE

INTUITION ACQUIRED BY OBSERVATION

Exploration of Chemical Spaces



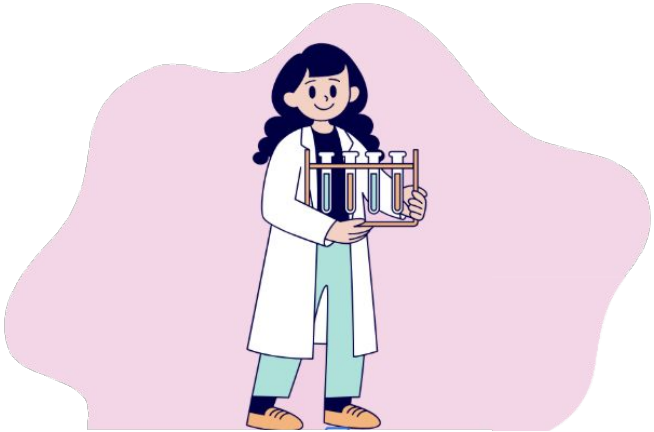


PREVIOUS KNOWLEDGE

INTUITION ACQUIRED BY OBSERVATION

Imagine new compounds / structure of interest.

Exploration of Chemical Spaces





PREVIOUS KNOWLEDGE

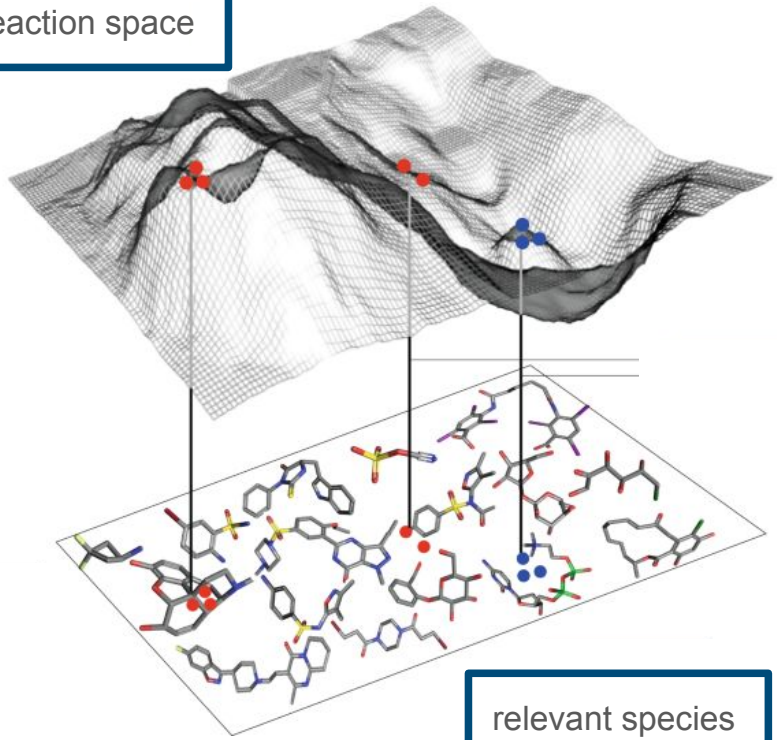
INTUITION ACQUIRED BY OBSERVATION

Imagine new compounds / structure of interest.

reaction discovery remains largely heuristic and human-driven

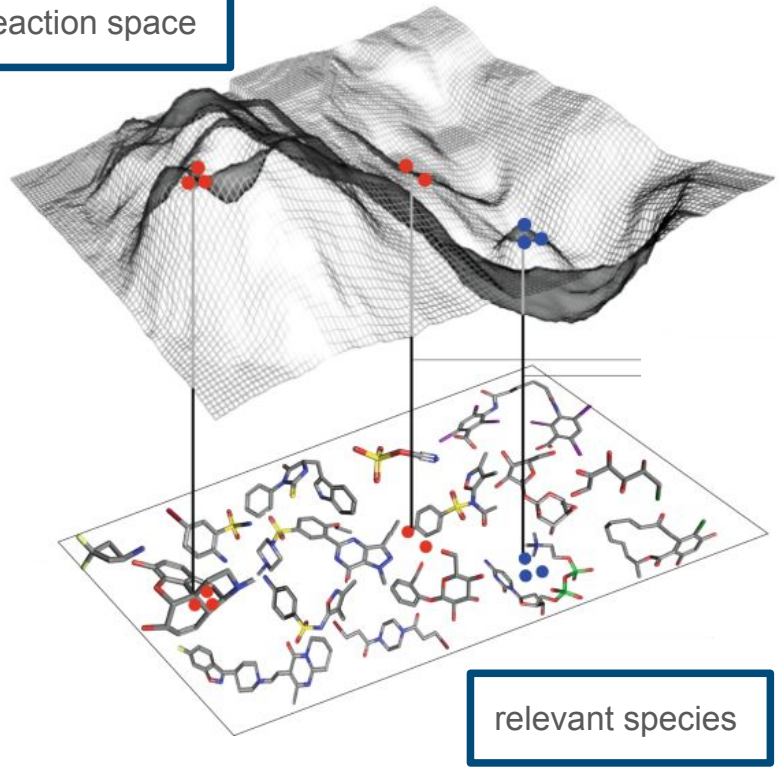
Exploration of Chemical Spaces

chemical reaction space



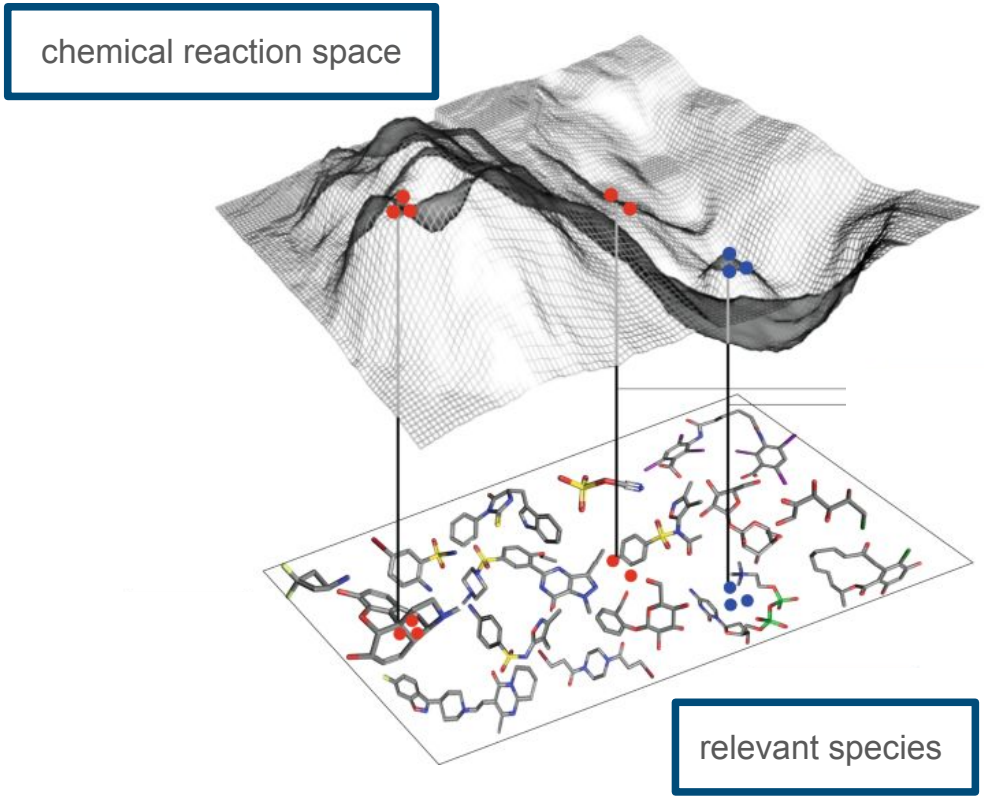
relevant species

Exploration of Chemical Spaces



Systematically exploring chemical space is limited by:

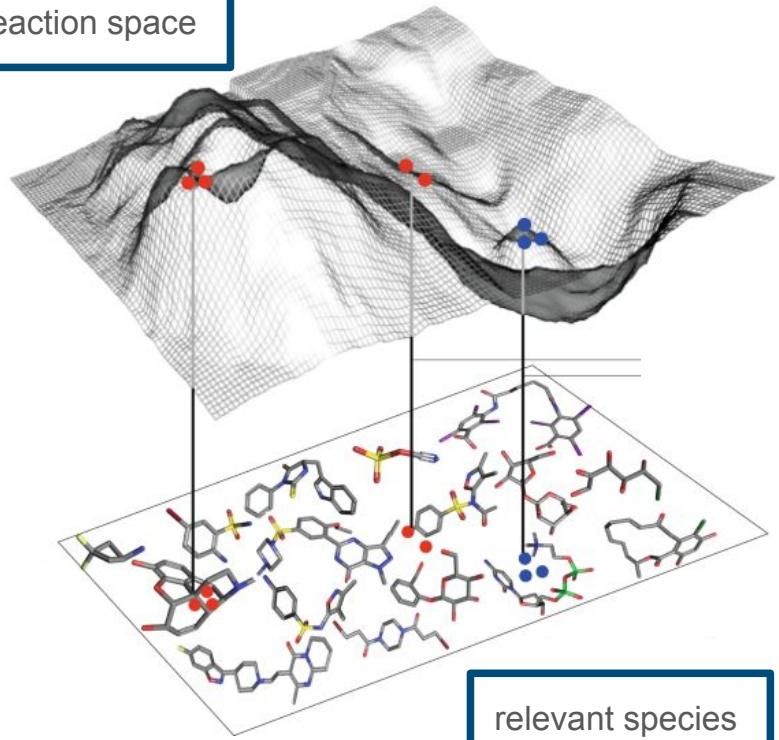
Exploration of Chemical Spaces



Systematically exploring chemical space is limited by:

- computational cost of quantum methods

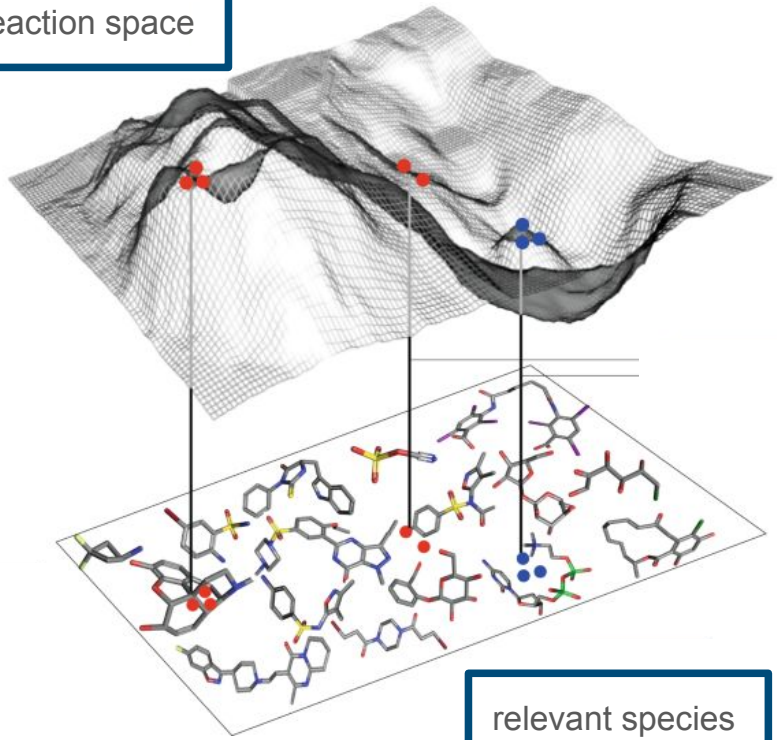
Exploration of Chemical Spaces



Systematically exploring chemical space is limited by:

- computational cost of quantum methods
- difficulty to find reaction paths (creation of large databases) as chemist input required

Exploration of Chemical Spaces



Systematically exploring chemical space is limited by:

- computational cost of quantum methods
- difficulty to find reaction paths (creation of large databases) as chemist input required
- ability to extract relevant information from large databases

Exploration of Chemical Spaces

Systematically exploring chemical space is limited by:



- computational cost of quantum methods

- difficulty to find reaction paths (creation of large databases) as chemist input required

- ability to extract relevant information from large databases

Exploration of Chemical Spaces

Systematically exploring chemical space is limited by:

- computational cost of quantum methods
- difficulty to find reaction paths (creation of large databases) as chemist input required
- ability to extract relevant information from large databases



partially solved with the use of MLIP

Exploration of Chemical Spaces

Systematically exploring chemical space is limited by:

- computational cost of quantum methods

partially solved with the use of MLIP



- difficulty to find reaction paths (creation of large databases) as chemist input required



- ability to extract relevant information from large databases

Exploration of Chemical Spaces

Systematically exploring chemical space is limited by:

- computational cost of quantum methods

partially solved with the use of MLIP

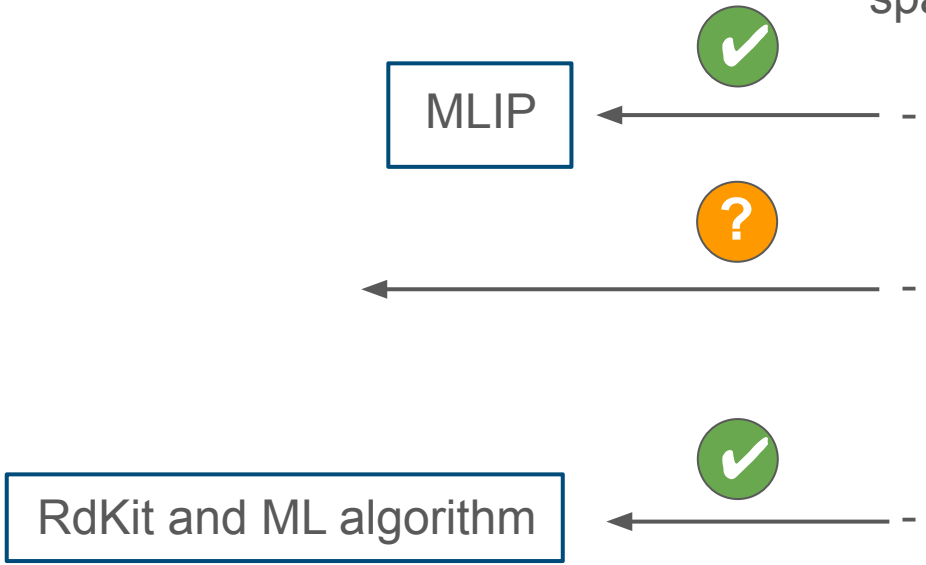
- difficulty to find reaction paths (creation of large databases) as chemist input required

- ability to extract relevant information from large databases

development of Python libraries such as RDKit and ML algorithms that enable overcoming these difficulties

Exploration of Chemical Spaces

Systematically exploring chemical space is limited by:



- computational cost of quantum methods
- difficulty to find reaction paths (creation of large databases) as chemist input required
- ability to extract relevant information from large databases

Exploration of Chemical Spaces

TS-tools project:
Automatic TS search localization program, boosted by reaction path search done using UMA MLIP and GPUs

MLIP ✔ - computational cost of quantum methods

? - difficulty to find reaction paths (creation of large databases) as chemist input required

RdKit and ML algorithm ✔ - ability to extract relevant information from large databases

Systematically exploring chemical space is limited by:

Exploration of Chemical Spaces - TS-tools

Received: 23 January 2024 | Revised: 8 March 2024 | Accepted: 20 March 2024
DOI: 10.1002/jcc.27374

SOFTWARE NOTE

Journal of COMPUTATIONAL CHEMISTRY WILEY

TS-tools: Rapid and automated localization of transition states based on a textual reaction SMILES input

Thijs Stuyver

Ecole Nationale Supérieure de Chimie de Paris, Université PSL, CNRS, Institute of Chemistry for Life and Health Sciences, Paris, France

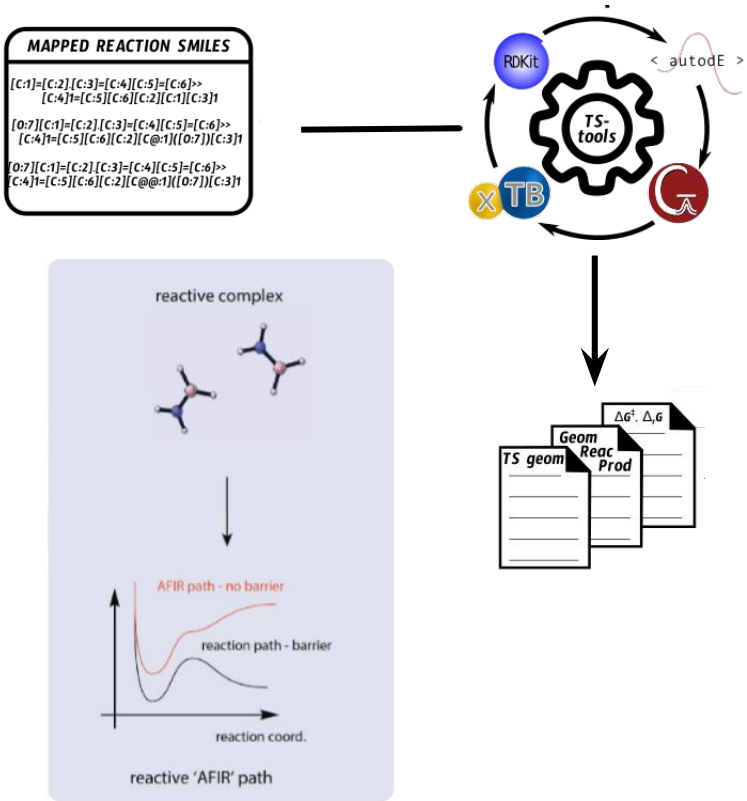
Correspondence
Thijs Stuyver, Ecole Nationale Supérieure de Chimie de Paris, Université PSL, CNRS, Institute of Chemistry for Life and Health Sciences, 75 005 Paris, France.
Email: thijs.stuyver@chimieparistech.psl.eu

Funding information
Agence Nationale de la Recherche

Abstract

Here, TS-tools is presented, a Python package facilitating the automated localization of transition states (TS) based on a textual reaction SMILES input. TS searches can either be performed at xTB or DFT level of theory, with the former yielding guesses at marginal computational cost, and the latter directly yielding accurate structures at greater expense. On a benchmarking dataset of mono- and bimolecular reactions, TS-tools reaches an excellent success rate of 95% already at xTB level of theory. For tri- and multimolecular reaction pathways - which are typically not benchmarked when developing new automated TS search approaches, yet are relevant for various types of reactivity, cf. solvent- and autocatalysis and enzymatic reactivity - TS-tools retains its ability to identify TS geometries, though a DFT treatment becomes essential in many cases. Throughout the presented applications, a particular emphasis is placed on solvation-induced mechanistic changes, another issue that received limited attention in the automated TS search literature so far.

KEYWORDS
automation, DFT, SMILES, *transition states*, xTB



Exploration of Chemical Spaces - TS-tools

Development of a workflow enabling quick identification of new candidates for the synthesis of self-healing polymers

ZINC NP
DOI: 10.71606/coconut.cnpc0052
76615 molecules
For NAtoms <16
98 dienes / 428 dienophiles
=> ~400.000 reactions

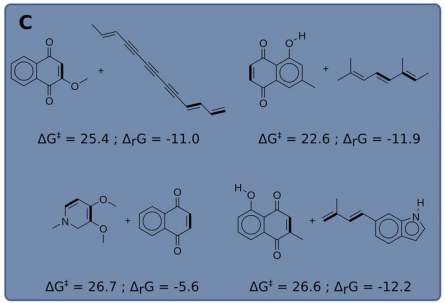
chemprop

$\Delta G^{\ddagger} < 30$
$-20 < \Delta_r G < 0$
4245 reactions
=
1277
diene/dienophile

DeepReaction

Obtaining DFT-quality profiles

32 reactions optimized at DFT level
28/32 respect the reversibility criteria



Exploration of Chemical Spaces - TS-tools

Development of a workflow enabling quick identification of new candidates for the synthesis of self-healing polymers

ZINC NP
DOI: 10.71606/coconut.cnpc0052
76615 molecules
For NAtoms <16
98 dienes / 428 die
=> ~400.000 rea

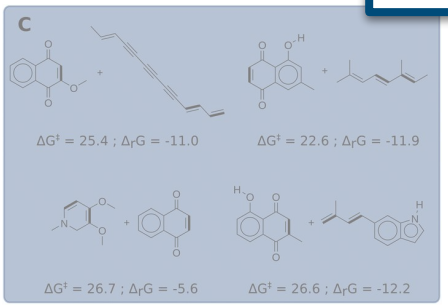
$\Delta G^{\ddagger} < 30$
$-20 < \Delta_r G < 0$
4245 reactions

DeepReaction

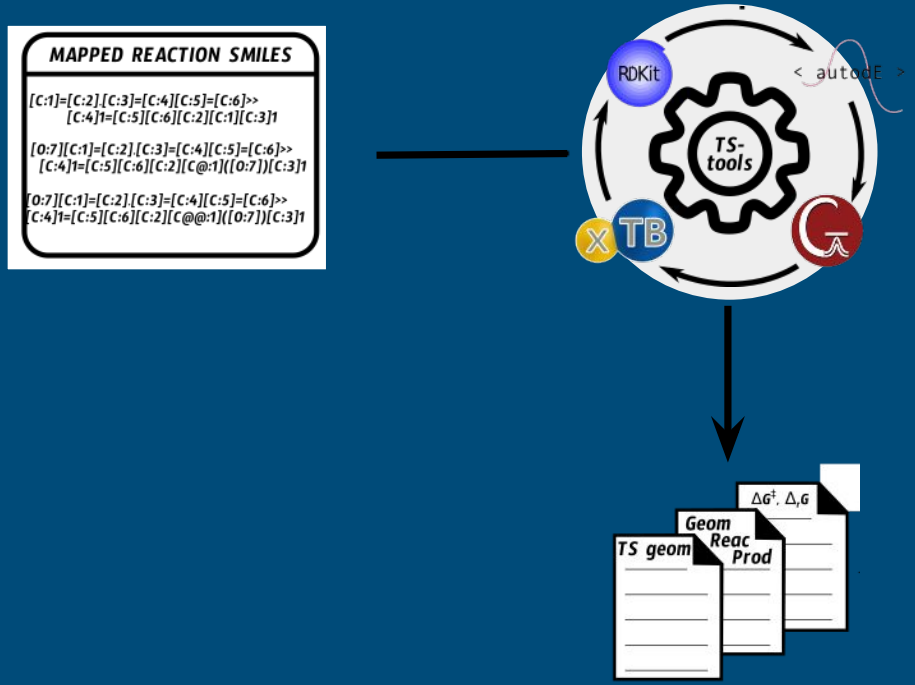
Applicable here because xTB correctly describes the Diels Alder reaction. Requires DFT validation anyway.

Obtaining DFT-quality profiles

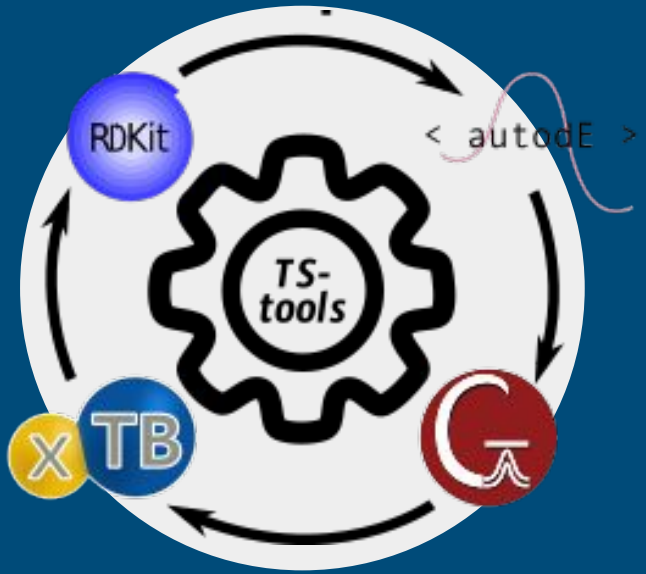
32 reactions optimized at DFT level
28/32 respect the reversibility criteria



Exploration of Chemical Spaces - TS-tools



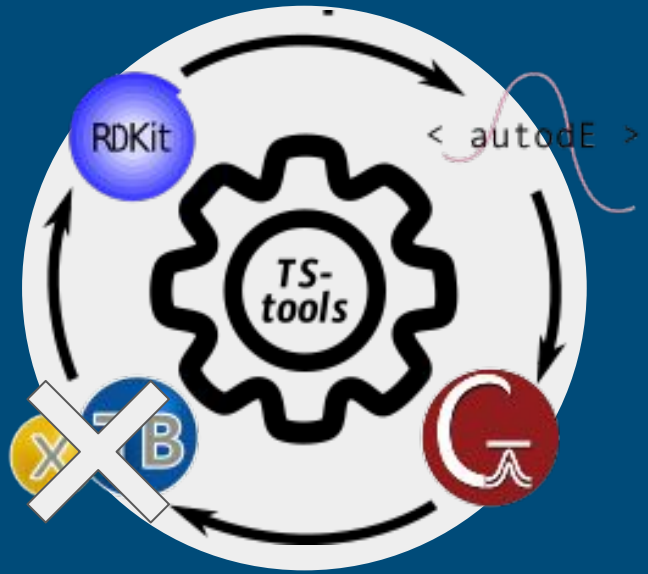
Exploration of Chemical Spaces - TS-tools



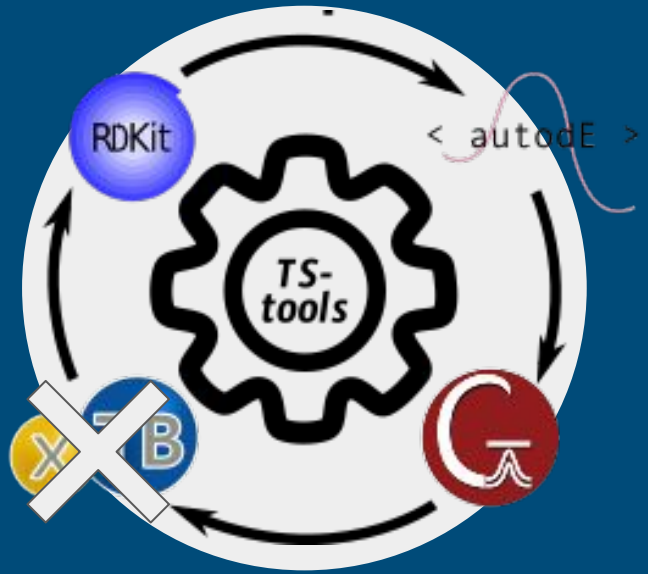
Exploration of Chemical Spaces - TS-tools



UMA - MLIP



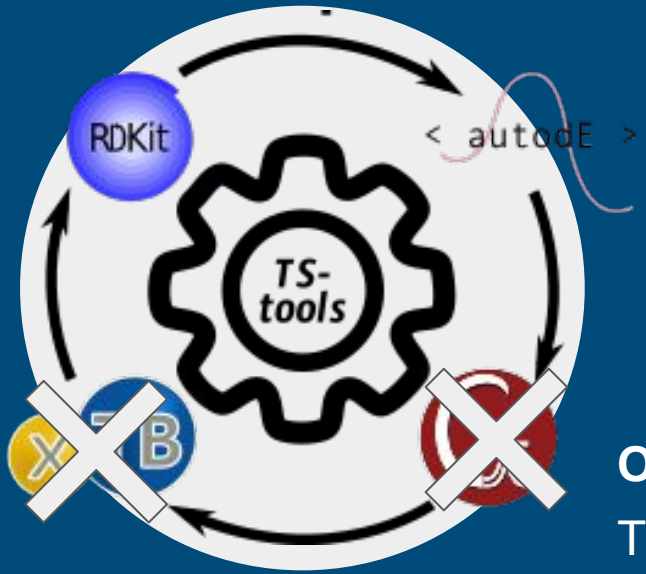
Exploration of Chemical Spaces - TS-tools



UMA - MLIP

\+ AFIR search run with batches on GPU

Exploration of Chemical Spaces - TS-tools



UMA - MLIP

\+ AFIR search run with batches on GPU

ORCA + UMA wrapper

TS guess optimization + IRC

Exploration of Chemical Spaces - TS-tools

Two-block code - Hybrid approach needed due to lack of GPU implementation for ORCA, and that the interface between ORCA and the MLIP is based on a SERVER/CLIENT architecture.

Exploration of Chemical Spaces - TS-tools

Two-block code - Hybrid approach needed due to lack of GPU implementation for ORCA, and that the interface between ORCA and the MLIP is based on a SERVER/CLIENT architecture.

GPU block

AFIR formalism

Generation of reaction path to generate TS guesses

Reactions run in batches for each artificial reaction factors

Exploration of Chemical Spaces - TS-tools

Two-block code - Hybrid approach needed due to lack of GPU implementation for ORCA, and that the interface between ORCA and the MLIP is based on a SERVER/CLIENT architecture.

GPU block

AFIR formalism

Generation of reaction path to generate TS guesses

Reactions run in batches for each artificial reaction factors

CPU block

TS optimization and verification through IRC computations

UMA model load on a CPU node (server). Every time ORCA (client) generates a new guess structure it ask the server to determine the electronic energy of this new structure.

Exploration of Chemical Spaces - TS-tools

Two-block code - Hybrid approach needed due to lack of GPU implementation

```
INFO:uma-server:model loaded in 2.9s
INFO:uma-server:shared 149 params + 82 buffers (586 MB) via /dev/shm
[Master] 1 worker(s) on cpu, model /home/ferrerm/TS-tools-NNP_new/ORCA_wrapper_UMA-main/models/uma-s-1p1.pt, core budget 8, port 20261
INFO:uma-server:ready on 127.0.0.1:20261 (core budget: 8)
INFO:uma-server:[W0] OK  0.827s (8T)  /home/ferrerm/TS-tools-NNP_new/test_wrapper/work_dir/orca_dir_3/exp_R1_144_TS_0_EXT.xyz
INFO:uma-server:[W0] OK  0.714s (8T)  /home/ferrerm/TS-tools-NNP_new/test_wrapper/work_dir/orca_dir_3/CC_R1_118_3290_TS_0_EXT.xyz
INFO:uma-server:[W0] OK  0.584s (8T)  /home/ferrerm/TS-tools-NNP_new/test_wrapper/work_dir/orca_dir_3/CC_R1_118_3290_TS_3_EXT.xyz
INFO:uma-server:[W0] OK  1.013s (8T)  /home/ferrerm/TS-tools-NNP_new/test_wrapper/work_dir/orca_dir_3/exp_R1_144_TS_0_EXT.xyz
INFO:uma-server:[W0] OK  0.652s (8T)  /home/ferrerm/TS-tools-NNP_new/test_wrapper/work_dir/orca_dir_3/CC_R1_118_3290_TS_1_EXT.xyz
INFO:uma-server:[W0] OK  0.212s (8T)  /home/ferrerm/TS-tools-NNP_new/test_wrapper/work_dir/orca_dir_3/CC_R1_118_3290_TS_1_EXT.xyz
INFO:uma-server:[W0] OK  0.385s (8T)  /home/ferrerm/TS-tools-NNP_new/test_wrapper/work_dir/orca_dir_3/CC_R1_118_3290_TS_3_EXT.xyz
INFO:uma-server:[W0] OK  0.743s (8T)  /home/ferrerm/TS-tools-NNP_new/test_wrapper/work_dir/orca_dir_3/exp_R1_144_TS_0_EXT.xyz
INFO:uma-server:[W0] OK  0.397s (8T)  /home/ferrerm/TS-tools-NNP_new/test_wrapper/work_dir/orca_dir_3/CC_R1_118_3290_TS_0_EXT.xyz
INFO:uma-server:[W0] OK  0.314s (8T)  /home/ferrerm/TS-tools-NNP_new/test_wrapper/work_dir/orca_dir_3/CC_R1_118_3290_TS_1_EXT.xyz
INFO:uma-server:[W0] OK  0.333s (8T)  /home/ferrerm/TS-tools-NNP_new/test_wrapper/work_dir/orca_dir_3/CC_R1_118_3290_TS_3_EXT.xyz
INFO:uma-server:[W0] OK  0.244s (8T)  /home/ferrerm/TS-tools-NNP_new/test_wrapper/work_dir/orca_dir_3/CC_R1_118_3290_TS_1_EXT.xyz
INFO:uma-server:[W0] OK  0.273s (8T)  /home/ferrerm/TS-tools-NNP_new/test_wrapper/work_dir/orca_dir_3/CC_R1_118_3290_TS_2_EXT.xyz
INFO:uma-server:[W0] OK  0.295s (8T)  /home/ferrerm/TS-tools-NNP_new/test_wrapper/work_dir/orca_dir_3/CC_R1_118_3290_TS_3_EXT.xyz
INFO:uma-server:[W0] OK  0.208s (8T)  /home/ferrerm/TS-tools-NNP_new/test_wrapper/work_dir/orca_dir_3/CC_R1_118_3290_TS_3_EXT.xyz
INFO:uma-server:[W0] OK  0.266s (8T)  /home/ferrerm/TS-tools-NNP_new/test_wrapper/work_dir/orca_dir_3/CC_R1_118_3290_TS_2_EXT.xyz
INFO:uma-server:[W0] OK  0.384s (8T)  /home/ferrerm/TS-tools-NNP_new/test_wrapper/work_dir/orca_dir_3/CC_R1_118_3290_TS_1_EXT.xyz
INFO:uma-server:[W0] OK  0.864s (8T)  /home/ferrerm/TS-tools-NNP_new/test_wrapper/work_dir/orca_dir_3/exp_R1_144_TS_0_EXT.xyz
INFO:uma-server:[W0] OK  0.958s (8T)  /home/ferrerm/TS-tools-NNP_new/test_wrapper/work_dir/orca_dir_3/CC_R1_118_3290_TS_3_EXT.xyz
INFO:uma-server:[W0] OK  1.045s (8T)  /home/ferrerm/TS-tools-NNP_new/test_wrapper/work_dir/orca_dir_3/CC_R1_118_3290_TS_2_EXT.xyz
INFO:uma-server:[W0] OK  0.598s (8T)  /home/ferrerm/TS-tools-NNP_new/test_wrapper/work_dir/orca_dir_3/CC_R1_118_3290_TS_1_EXT.xyz
INFO:uma-server:[W0] OK  0.659s (8T)  /home/ferrerm/TS-tools-NNP_new/test_wrapper/work_dir/orca_dir_3/CC_R1_118_3290_TS_0_EXT.xyz
INFO:uma-server:[W0] OK  0.251s (8T)  /home/ferrerm/TS-tools-NNP_new/test_wrapper/work_dir/orca_dir_3/CC_R1_118_3290_TS_2_EXT.xyz
```

structure.

Exploration of Chemical Spaces - TS-tools

GPU block (2 NVIDIA GPU A30 (256 Gb RAM) and 1 NVIDIA RTX6000 pro (512 Gb RAM))

```
Path generator Bath Creation
Starting generation of paths
Reaction R1: Stretched Optimization
Reaction R2: Stretched Optimization
Reaction R1: Reactive Optimization
Reaction R2: Reactive Optimization
Reaction R1: Stretched Optimization
Reaction R2: Stretched Optimization
Reaction R1: Reactive Optimization
Reaction R2: Reactive Optimization
Reaction R1: Stretched Optimization
Reaction R2: Stretched Optimization
Reaction R1: Reaction Path Optimization
Reaction R2: Reactive Optimization
Reaction R1: Stretched Optimization
Reaction R2: Stretched Optimization
Reaction R1: Reaction Path Optimization
Reaction R2: Reactive Optimization
Reaction R1: Stretched Optimization
Reaction R2: Stretched Optimization
Reaction R1: Reaction Path Optimization
Reaction R2: Reaction Path Optimization
Reaction R1: Completed
Reaction R2: Stretched Optimization
Reaction R2: Reaction Path Optimization
Reaction R2: Stretched Optimization
Reaction R2: Reaction Path Optimization
Reaction R2: Stretched Optimization
```

```
Reaction R1: Stretched Optimization
Reaction R2: Stretched Optimization
Reaction R1: Reaction Path Optimization
Reaction R2: Reaction Path Optimization
Reaction R1: Completed
Reaction R2: Stretched Optimization
Reaction R2: Reaction Path Optimization
Reaction R2: Stretched Optimization
Reaction R2: Reaction Path Optimization
Reaction R2: Stretched Optimization
Reaction R2: Reaction Path Optimization
Reaction R2: Stretched Optimization
Reaction R2: Reaction Path Optimization
Reaction R2: Stretched Optimization
Reaction R2: Reaction Path Optimization
Reaction R2: Stretched Optimization
Reaction R2: Reaction Path Optimization
Reaction R2: Completed
Energies and Paths obtained
Energies written to energies.pkl
Paths written to paths.pkl
```

```
Devide used : cuda
Total number of paths to explore: 2
```

Time 183 seconds (~ 1.5 minutes per reaction/per reaction factor)

R1

[N+:1](=[B-:2](/[H:6])[H:7])(\[H:8])[H:9].[N+:3](=[B-:4](/[H:11])[H:12])(\[H:5])[H:10]**>>**[N+:1]([B-:2]([H:6])([H:7])[H:12])([B:4]([N:3]([H:5])[H:10])[H:11])([H:8])[H:9]

R2

[N+:1](=[B-:2](/[H:6])[H:7])(\[H:8])[H:9].[N+:3](=[B-:4](/[H:11])[H:12])(\[H:5])[H:10]**>>**[N+:1]([B-:2]([H:6])([H:7])[H:9])([B:4]([N:3]([H:5])[H:10])[H:11])([H:8])[H:12]

Exploration of Chemical Spaces - TS-tools

CPU block

```
Submitted array 0: 292816
All ORCA jobs finished
Submitted array 0: 292826
All ORCA jobs finished
Submitted array 0: 292827
All ORCA jobs finished
energies.pkl, paths.pkl removed
Successful reactions: ['R1']
Time taken: 183.62259531021118 seconds
```

~ 1.5 minutes per reaction, for TS optimization and IRC computations.

Rq: ORCA runs on only 1 CPU

Exploration of Chemical Spaces - TS-tools

CPU block

```
Submitted array 0: 292816
All ORCA jobs finished
Submitted array 0: 292826
All ORCA jobs finished
Submitted array 0: 292827
All ORCA jobs finished
energies.pkl, paths.pkl removed
Successful reactions: ['R1']
Time taken: 183.62259531021118 seconds
```

~ 1.5 minutes per reaction, for TS optimization and IRC computations.

Rq: ORCA runs on only 1 CPU

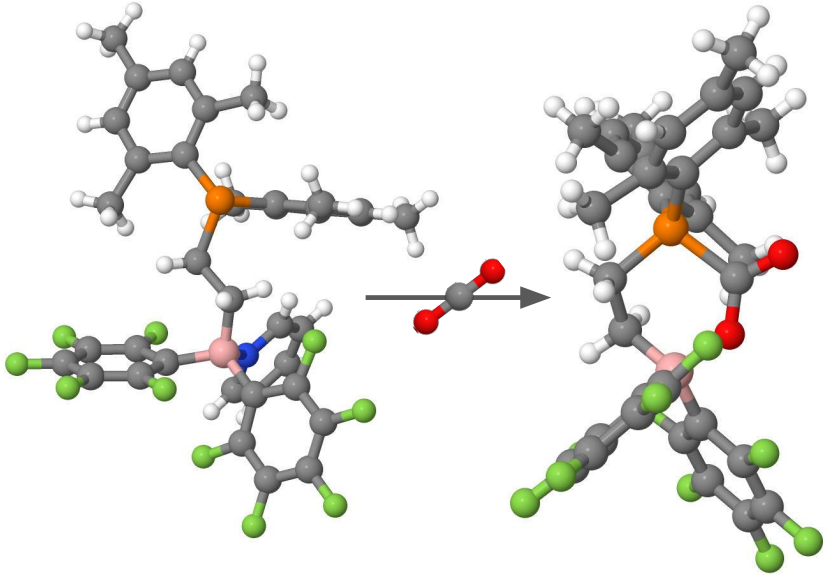
This hybrid approach is more efficient than running fully on GPU. Python libraries for TS optimization are not yet optimized to run batches on GPUs, so loss of efficiency.

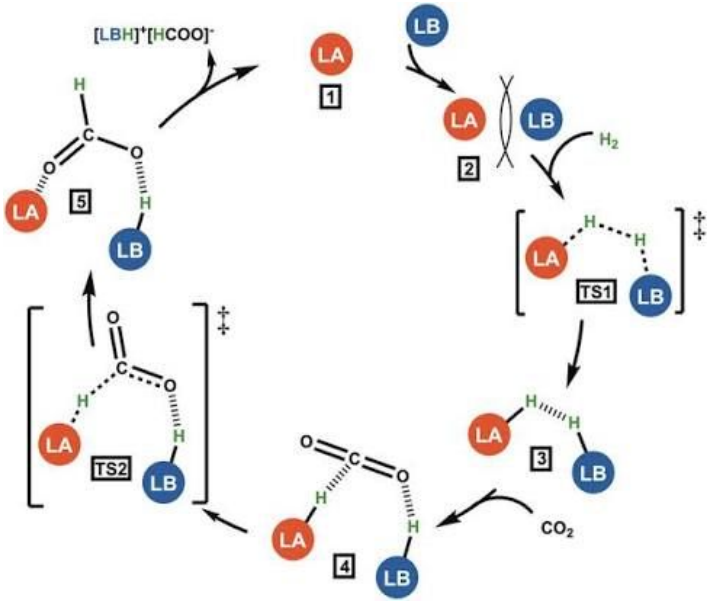
On the same two reactions
GPU total time = 1539 seconds.
Hybrid = 1023 seconds.

Rq: Of course the power of this approach will be more significant in the case of systems not tractable with DFT, and running large batches.

Exploration of Chemical Spaces - TS-tools

FLP case (large systems, more than 20 atoms, with heavy atoms)
In this case, DFT optimization of a minimum can take more than 24 hours on 8 CPUs and 32 GB of RAM.





from DOI:10.1002/anie.202208987

Exploration of Chemical Spaces - TS-tools

FLP case (large systems, more than 20 atoms, with heavy atoms)
In that case optimization at DFT, of a minima can take more than 24 hours on 8 CPUs and 32 GB of RAM.

Run a total of 338 reactions (**169 FLPs**)

Successful reactions : **136** (~40%)

Total time = **51 hours** (less than 3 times the time required to optimize a unique FLP minimum at DFT on 8 CPUs !!!)

Exploration of Chemical Spaces - TS-tools

FLP case (large systems, more than 20 atoms, with heavy atoms)
In that case optimization at DFT, of a minima can take more than 24 hours on 8 CPUs and 32 GB of RAM.

Run a total of 338 reactions (169 FLPs)
Successful reactions : 136 (~40%)
Total time = 51 hours (less than 3 times the time required to optimize a unique FLP minimum at DFT on 8 CPUs !!!)

We discovered that some FLPs were following **a 3-step mechanism** rather than a 2-step one.
If we consider only the H_2 capture: **120/169 successful reactions (~70%)**
Considering the flexibility of the system, and the absence of conformational search, we can consider that it is an interesting result.

Exploration of Chemical Spaces - TS-tools

FLP case (large systems, more than 20 atoms, with heavy atoms)
In that case optimization at DFT, of a minima can take more than 24 hours on 8 CPUs and 32 GB of RAM.

Run a total of 338 reactions (169 FLPs)

When considering large systems, non really tractable with DFT, the use of an hybrid approach on GPU enables significantly faster reaction discovery.
It is still necessary to optimize the TS-tools program to better deal with large and flexible systems, but the first results are really promising.

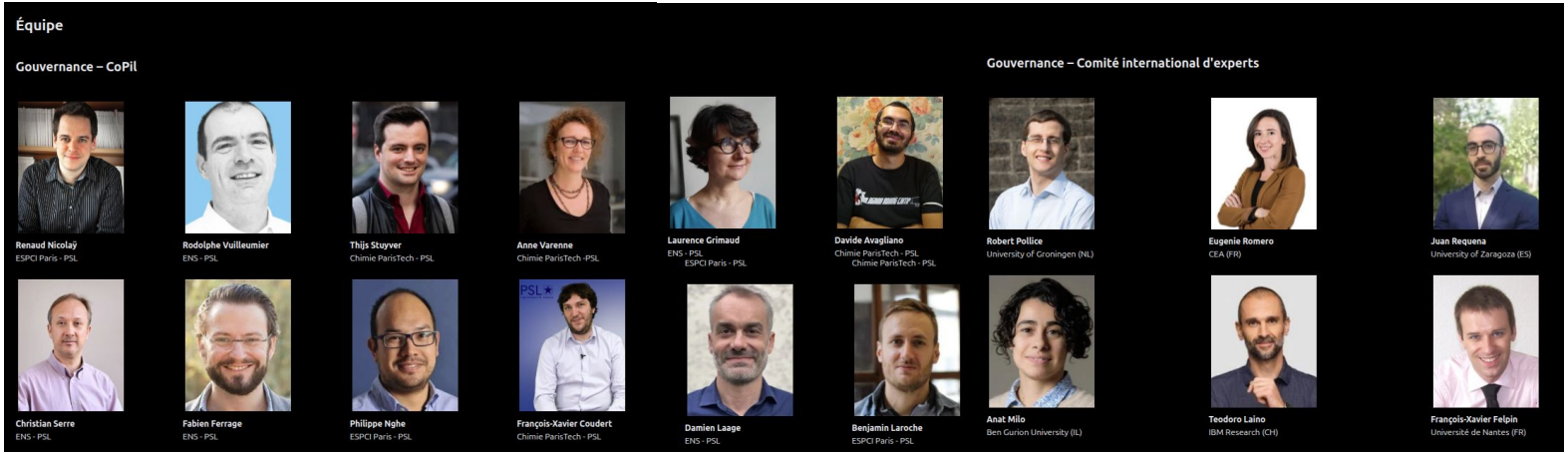
step one.
If we consider only the H_2 capture: 120/169 successful reactions (~70%)
Considering the flexibility of the system, and the absence of conformational search, we can consider that it is a interesting result.

Part of a bigger project

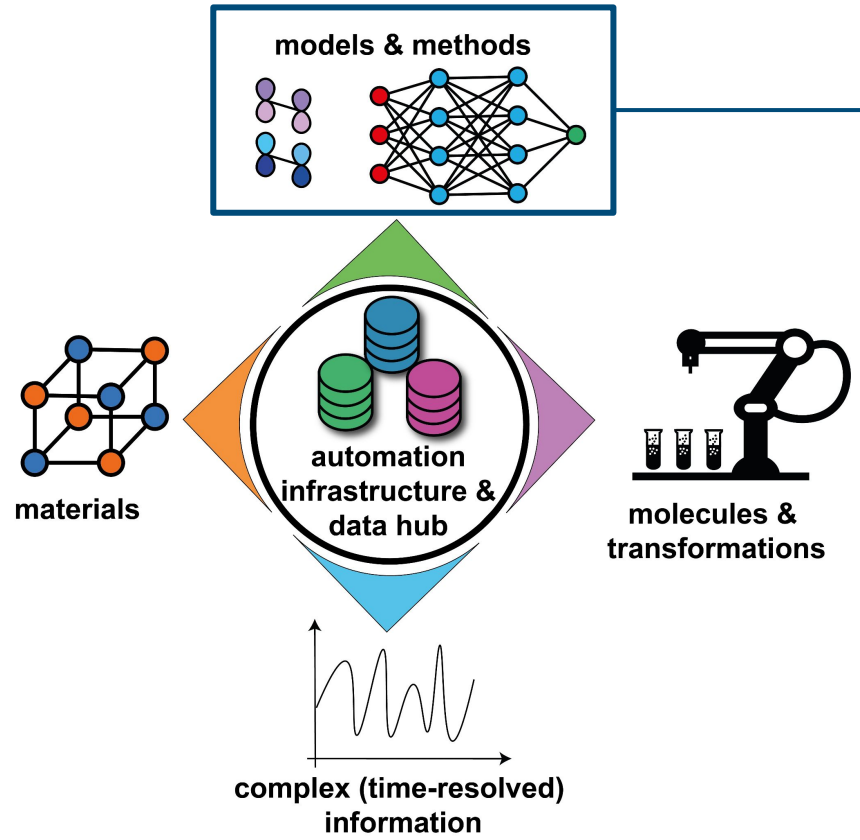


Part of a bigger project





Part of a bigger project

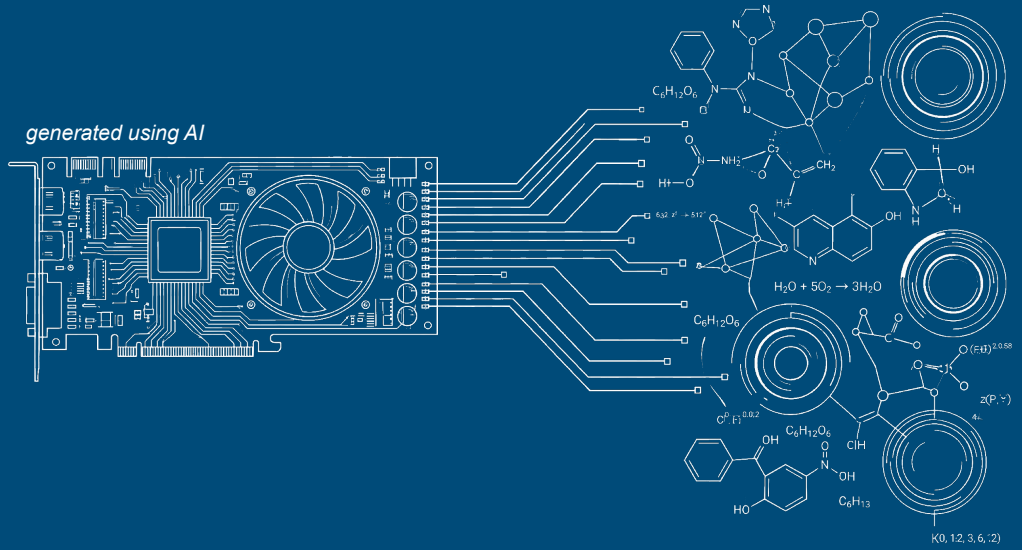


Development of new models and methods using ML and AI to resolve chemistry problems: automation platform, Bayesian optimization to optimize reactivity…

Applications that benefit strongly from GPU usage.

October 28th, 2026

Accelerating Molecular Discovery through GPU Computing



Thank you for your attention

Chimie Paris ParisTech | PSL | | FRANCE 2030 | PR[AI]RIE Paris School