XACS
Xiamen Atomistic Computing Suite
XACScloud.com

MLatom.com



Transition state search and analysis

Pavlo O. Dral

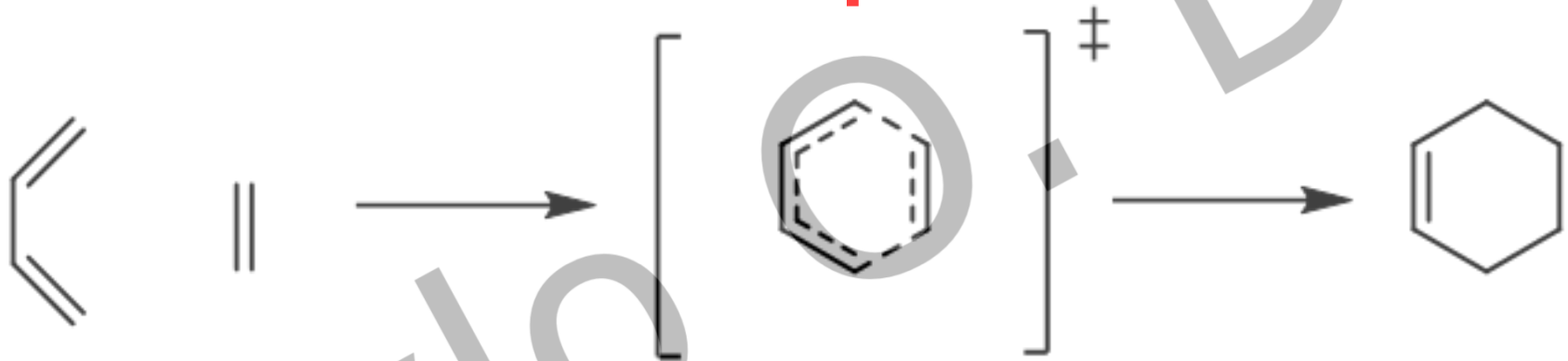
Xiamen University, P.R. China

Visiting Professor in
Nicolaus Copernicus University, Poland

5 July 2024

Pavlo O. Dral

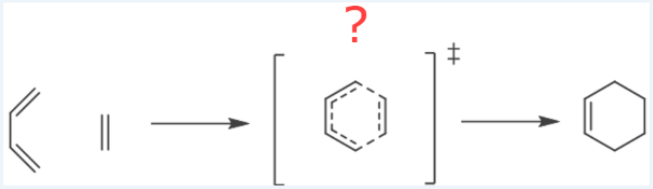
dr-dral.com



Pavlo O. Dral

Example ML-TS-1.

Try to find a TS for the Diels–Alder reaction between ethylene and butadiene:



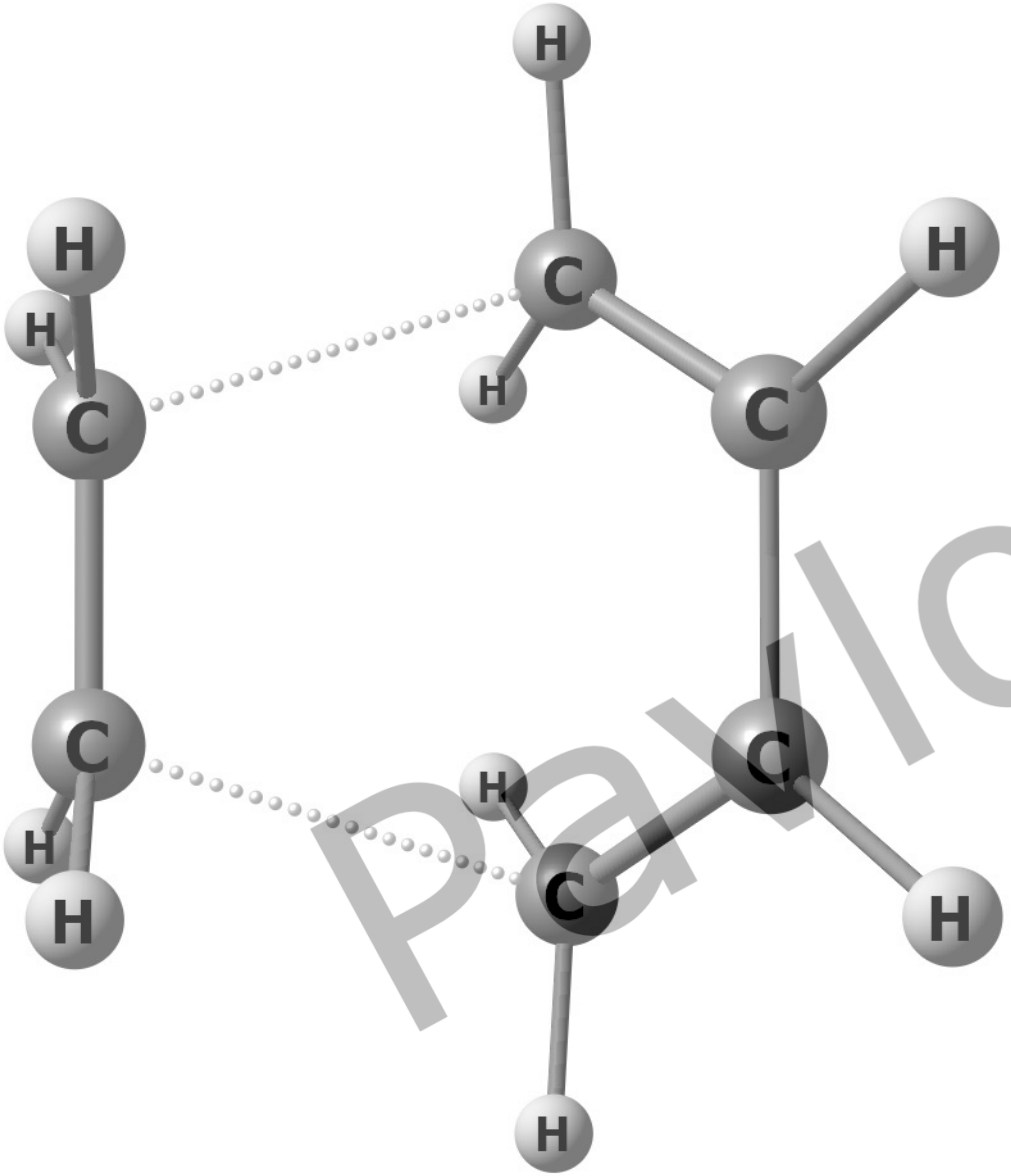
The most difficult part would be setting up a reasonable guess for the TS geometry used as an initial guess for TS optimization in MLatom. To simplify your life, we provide it.

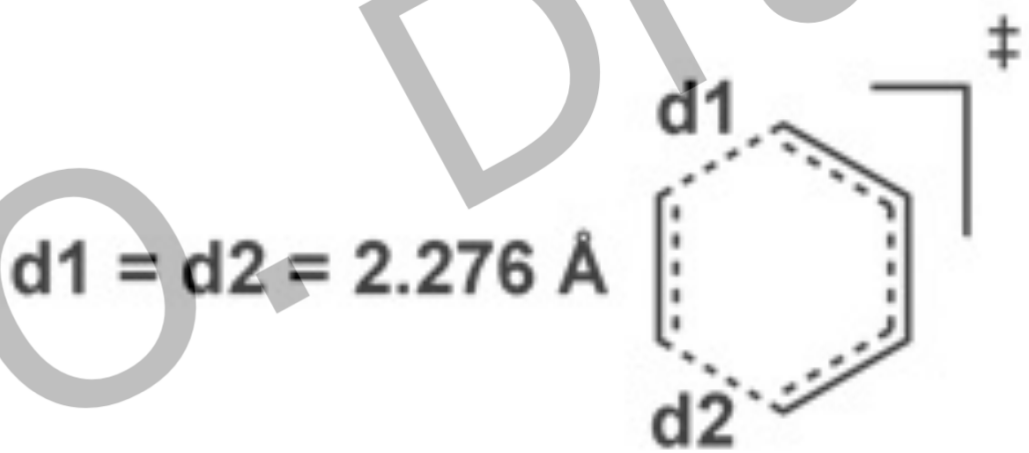
An input file:

```
uaiqm
ts
xyzfile='16

C          0.48462430     -0.55755495      1.43729151
C          0.48462430     -0.55755495     -1.43729151
C         -0.27595797     -1.44977527      0.70359025
C         -0.27595797     -1.44977527     -0.70359025
C         -0.27595797      1.45086377      0.69299925
C         -0.27595797      1.45086377     -0.69299925
H          0.37292526     -0.50748993      2.51767690
H          1.44526264     -0.21636383      1.06867438
H          0.37292526     -0.50748993     -2.51767690
H          1.44526264     -0.21636383     -1.06867438
H         -1.05536225     -2.01444047      1.21328943
H         -1.05536225     -2.01444047     -1.21328943
H          0.51071931      1.96707995      1.23581344
H         -1.20625330      1.32768072      1.23591744
H          0.51071931      1.96707995     -1.23581344
H         -1.20625330      1.32768072     -1.23591744
'
optxyz=da_ts_opt.xyz
```

Pavlo O. Dral

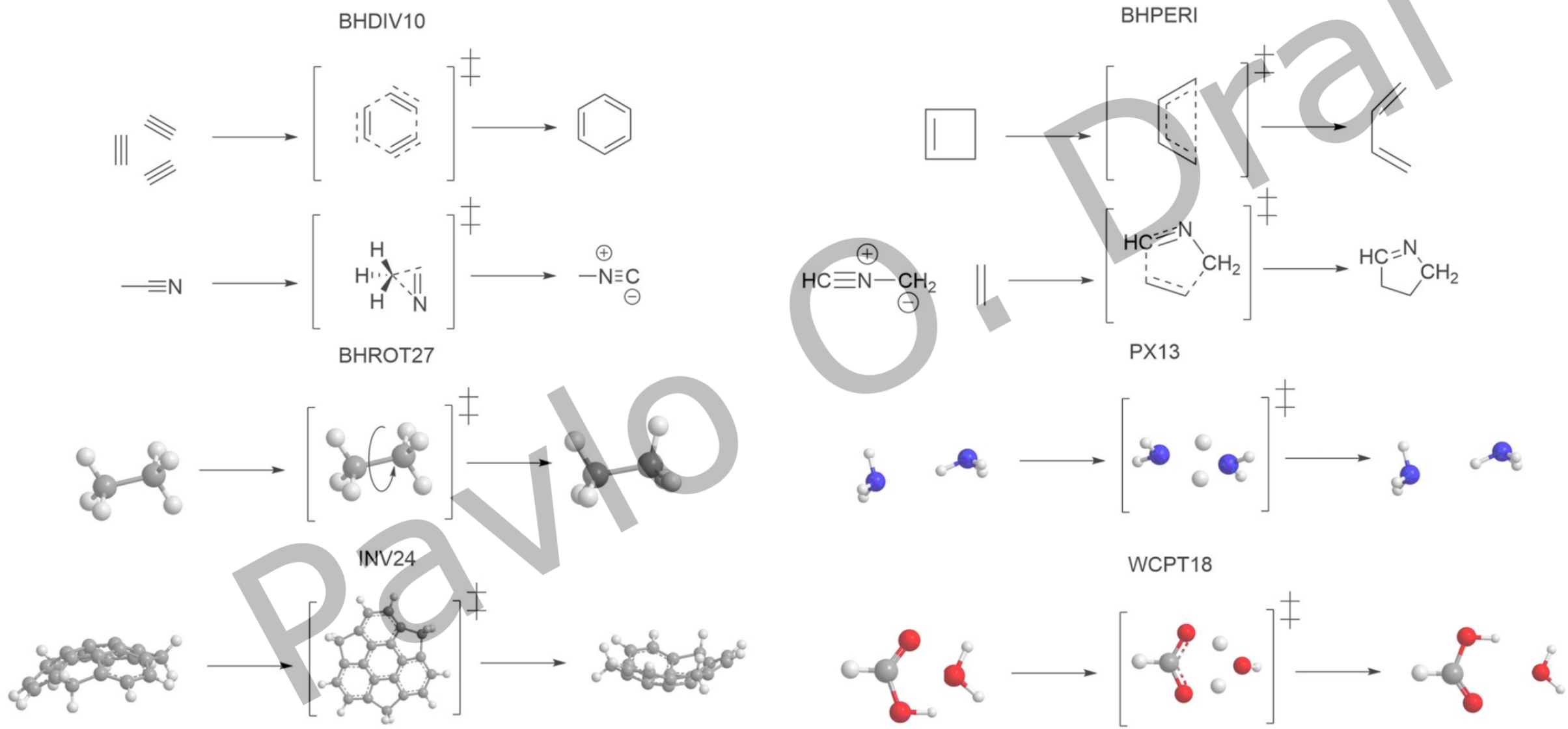




Depends on a level, but something like this (here B3LYP-D4/6-31G*)

Pavlo O. Dral

Reaction barrier with AIQM1



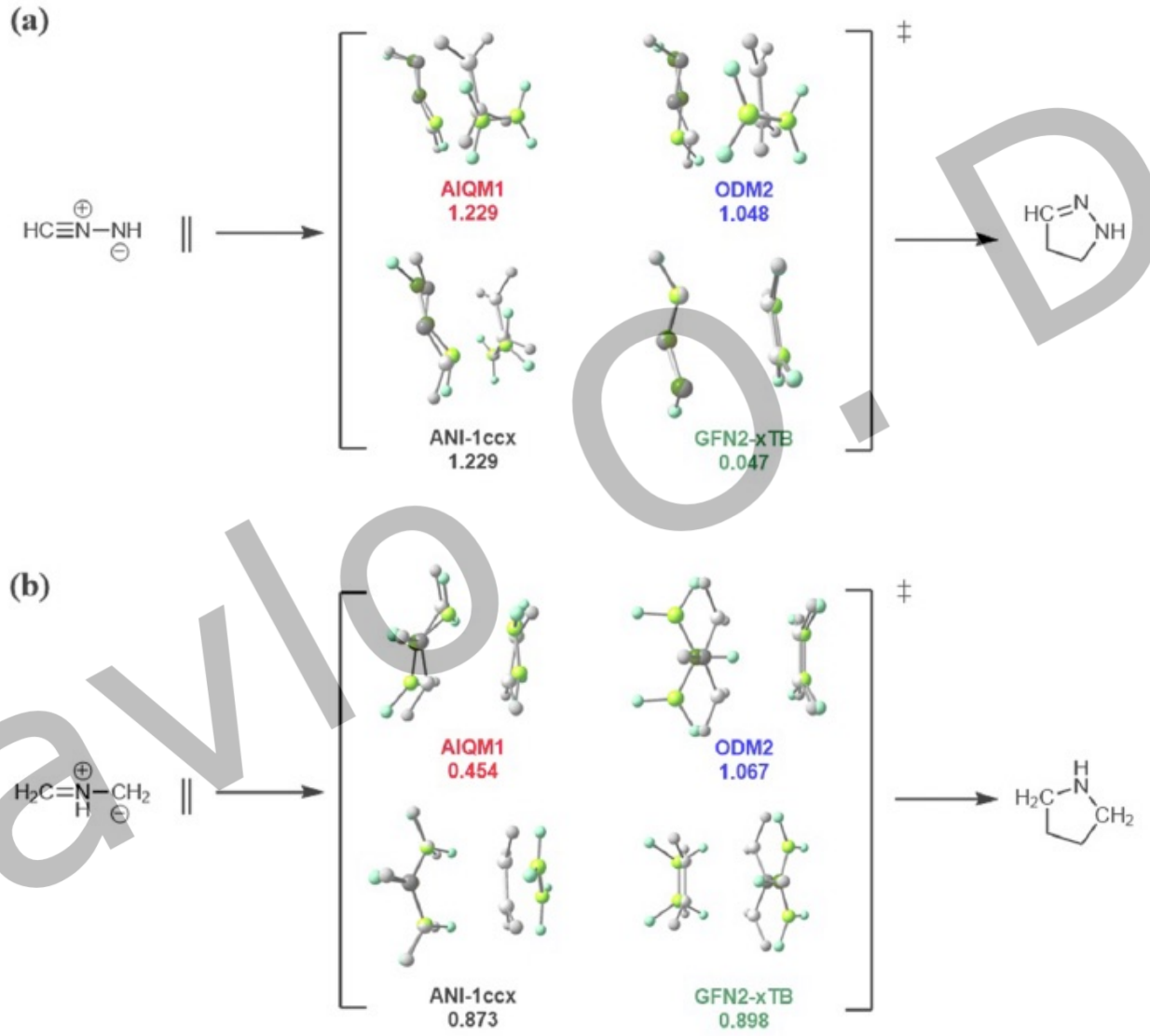
Pavlo O. Dral

Y. Chen, Y. Ou, P. Zheng, Y. Huang, F. Ge, P. O. Dral. *J. Chem. Phys.* **2023,** *158*, 074103

subsets	AIQM1	UAIQM_gfn2xtb	DSD-BLYP-D4/ 6-31G*	B3LYP-D4/ 6-31G*
BHPERI	10.90	2.54	4.07	4.02
BHDIV10	6.90	6.97	4.21	4.22
INV24	3.39	1.07	1.27	1.68
BHROT27	0.47	0.43	0.66	0.75
PX13	6.88	5.99	2.59	8.34
WCPT18	4.30	2.85	4.17	6.49
WTMAD2	5.45	2.24	2.84	3.53

Pavlo O. Dral

Y. Chen, Y.-F. Hou, O. Isayev, P. O. Dral. *Universal and Updatable Artificial Intelligence-Enhanced Quantum Chemical Foundational Models*. **2024**, *submitted*. https://doi.org/10.26434/chemrxiv-2024-604wb.


(a)
AIQM1
1.229
ODM2
1.048
ANI-1ccx
1.229
GFN2-xTB
0.047
(b)
AIQM1
0.454
ODM2
1.067
ANI-1ccx
0.873
GFN2-xTB
0.898

Pavlo O. Dral

XACS
Xiamen Atomistic Computing Suite
XACScloud.com

nature chemistry

MLatom.com

Article

https://doi.org/10.1038/s41557-023-01427-3

Exploring the frontiers of condensed-phase chemistry with a general reactive machine learning potential

ANI-1xnr???

Received: 13 April 2023

Accepted: 12 December 2023

Published online: 07 March 2024

Check for updates

Shuhao Zhang[1,2], Małgorzata Z. Makoś[3,4], Ryan B. Jadrich[2,5], Elfi Kraka[3], Kipton Barros[2,5], Benjamin T. Nebgen[2], Sergei Tretiak[2,5], Olexandr Isayev[1], Nicholas Lubbers[4] ✉, Richard A. Messerly[2] ✉ & Justin S. Smith[2,6] ✉

Atomistic simulation has a broad range of applications from drug design to materials discovery. Machine learning interatomic potentials (MLIPs) have become an efficient alternative to computationally expensive ab initio simulations. For this reason, chemistry and materials science would greatly benefit from a general reactive MLIP, that is, an MLIP that is applicable to a broad range of reactive chemistry without the need for refitting. Here we develop a general reactive MLIP (ANI-1xnr) through automated sampling of condensed-phase reactions. ANI-1xnr is then applied to study five distinct systems: carbon solid-phase nucleation, graphene ring formation from acetylene, biofuel additives, combustion of methane and the spontaneous formation of glycine from early earth small molecules. In all studies, ANI-1xnr closely matches experiment (when available) and/or previous studies using traditional model chemistry methods. As such, ANI-1xnr proves to be a highly general reactive MLIP for C, H, N and O elements in the condensed phase, enabling high-throughput in silico reactive chemistry experimentation.

Pavlo O. Dral

dr-dral.com

Example ML-TS-2.

For the structure(s) you found in the previous task, calculate frequencies:

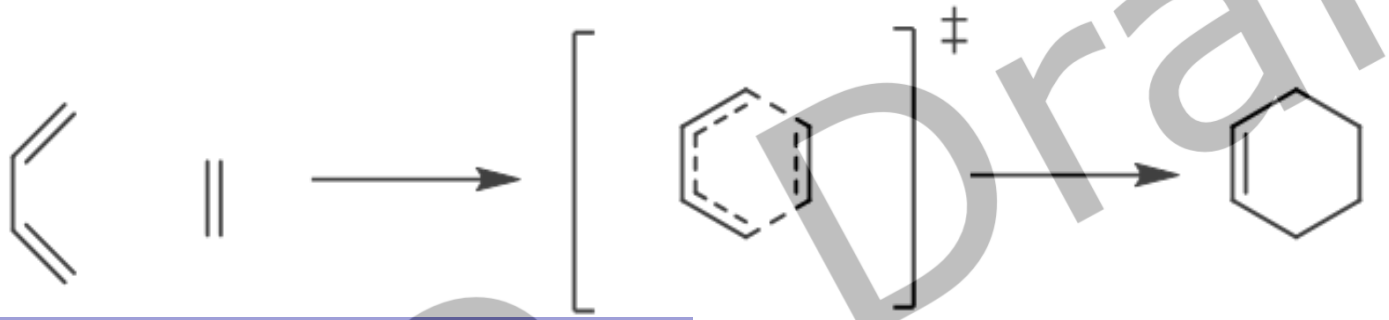
```
freq
uaiqm
xyzfile='16

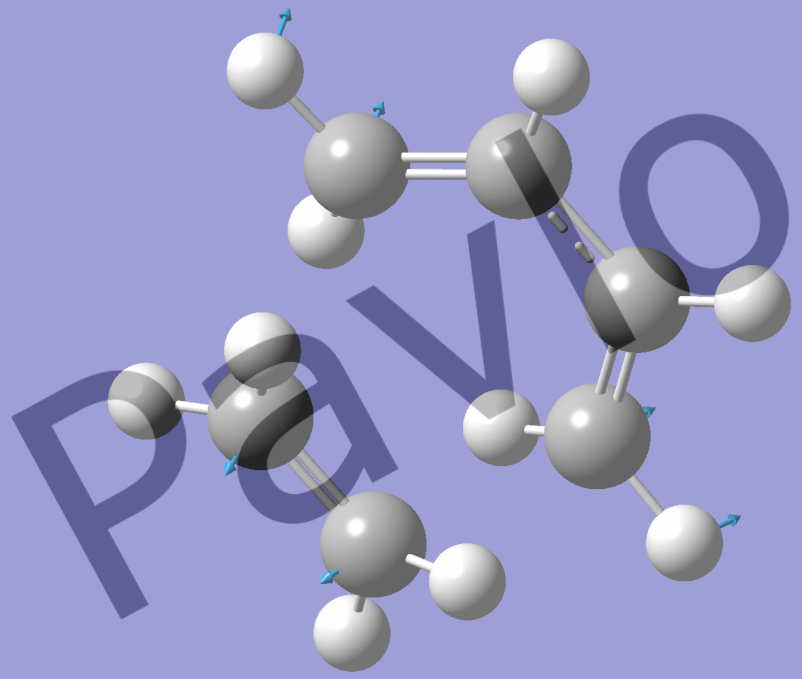
C             0.4984885582383          -0.5338087844381           1.4145928440825
C             0.4960692097735          -0.5384413336966          -1.4181301495364
C            -0.2896532359000          -1.4128506592682           0.7068921658160
C            -0.2905515192621          -1.4145888148315          -0.7051528877014
C            -0.2588083017453           1.4282744245852           0.6874347906976
C            -0.2574941435174           1.4288896351987          -0.6919434751381
H             0.3947104660660          -0.4397347928006           2.4878894665433
H             1.4467820702090          -0.2162549498538           1.0088131803715
H             0.3876593222716          -0.4479872886549          -2.4913060754311
H             1.4458716855847          -0.2184181276041          -1.0177581601047
H            -1.0783691986254          -1.9558675609816           1.2150705221731
H            -1.0799202791492          -1.9588818640283          -1.2105952858916
H             0.5409477501467           1.9087387905544           1.2339324532537
H            -1.1792980399036           1.2603193266422           1.2295776489298
H             0.5433702777959           1.9074275278623          -1.2394354082521
H            -1.1774768154298           1.2605820031641          -1.2339285873837
'
```

Pavlo O. Dral

Note

If you wonder how to make such an animation: we used Gaussian to run frequency calculations with our model, and the Gaussian output can be visualized with Gaussview or ChemCraft. We can't provide Gaussian on the cloud computing though. We will provide a work-around soon – please subscribe to our updates on WeChat!

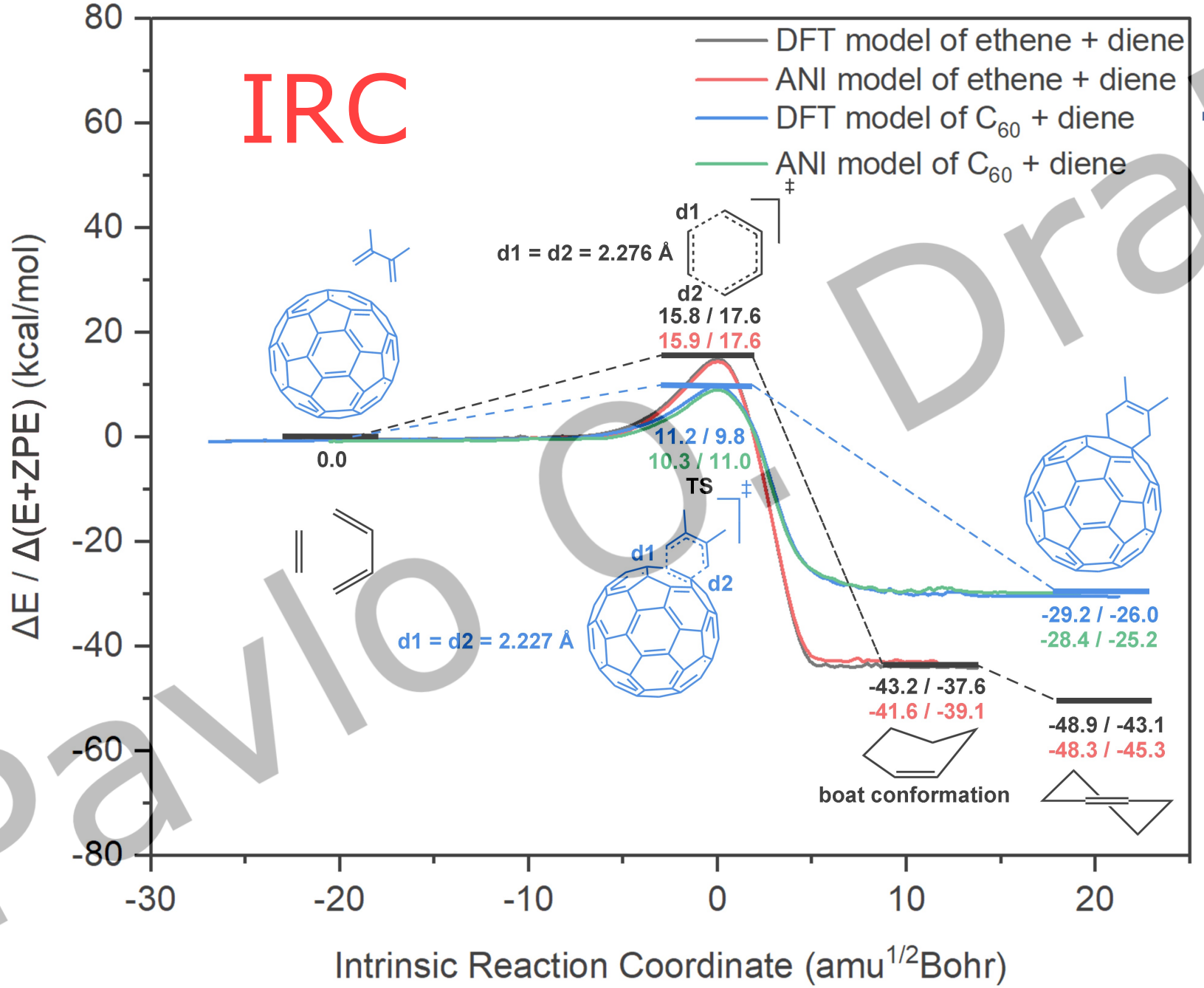




The output will contain one negative frequency:

Mode	Frequencies (cm^-1)	Reduced masses (AMU)	Force Constants (mDyne/A)
1	-599.9203	8.3844	-1.7779
2	55.2239	5.8098	0.0104
3	137.6350	2.0276	0.0226
4	259.0093	2.8785	0.1138
5	364.0193	3.9539	0.3087
6	375.8171	2.3977	0.1995
...			

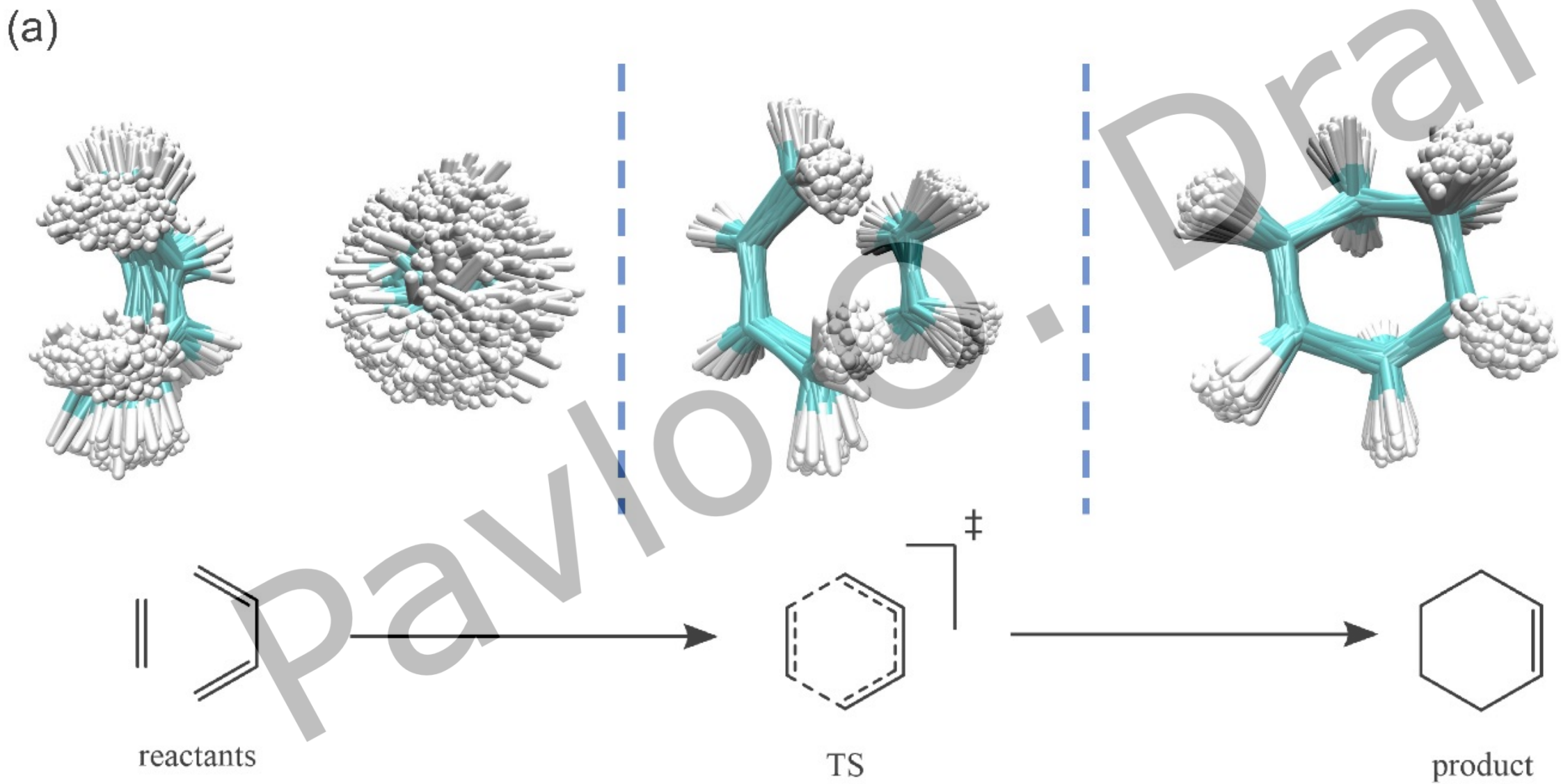
Pavlo O. Dral



Currently, not available on the cloud (needs Gaussian)

Pavlo O. Dral

Y.-F. Hou, Q. Zhang, P. O. Dral. https://doi.org/10.26434/chemrxiv-2024-hwsl4.



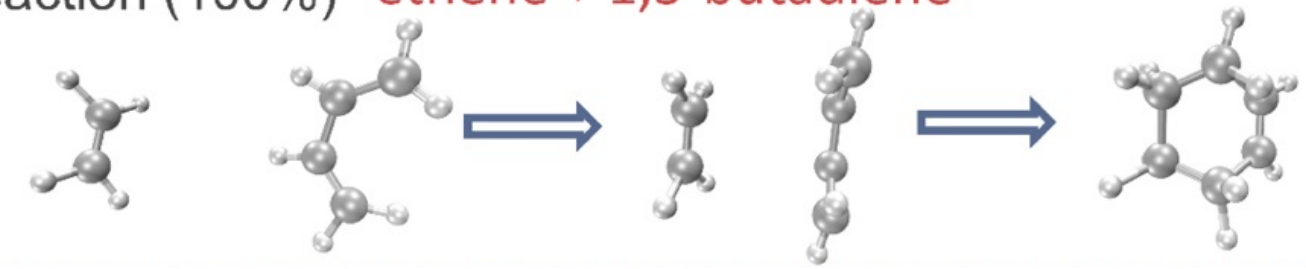
Pavlo O. Dral

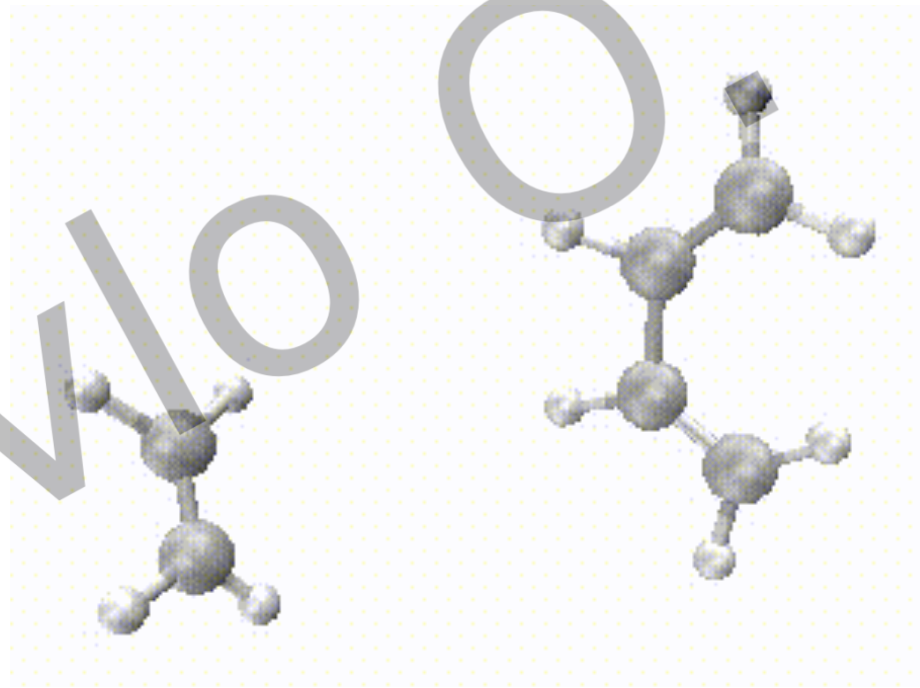
Y.-F. Hou, L. Zhang, Q. Zhang, F. Ge, P. O. Dral. *arXiv:2404.11811*.

Time-resolved mechanisms

~3000 training points with ANI potential

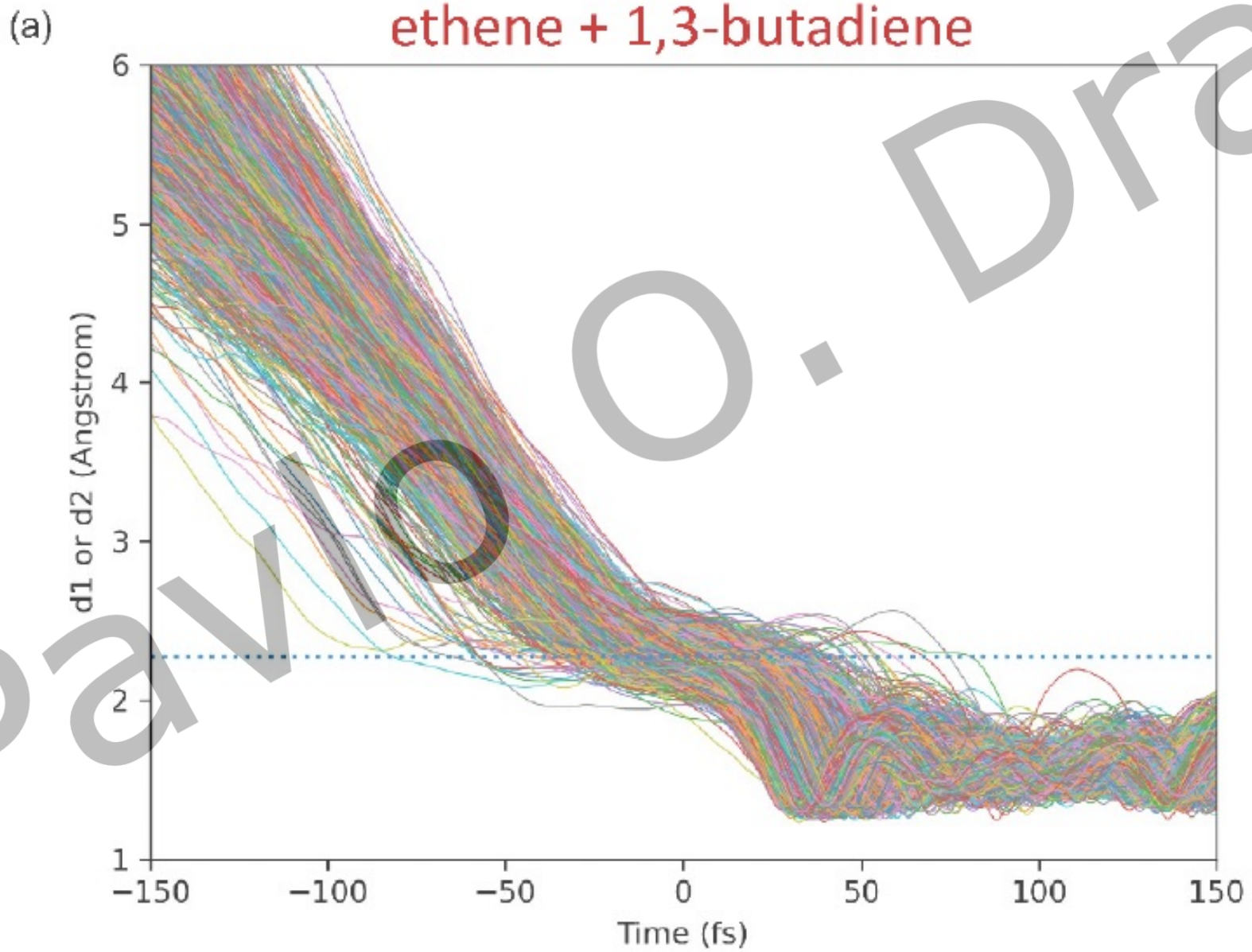
Direct reaction (100%) ethene + 1,3-butadiene



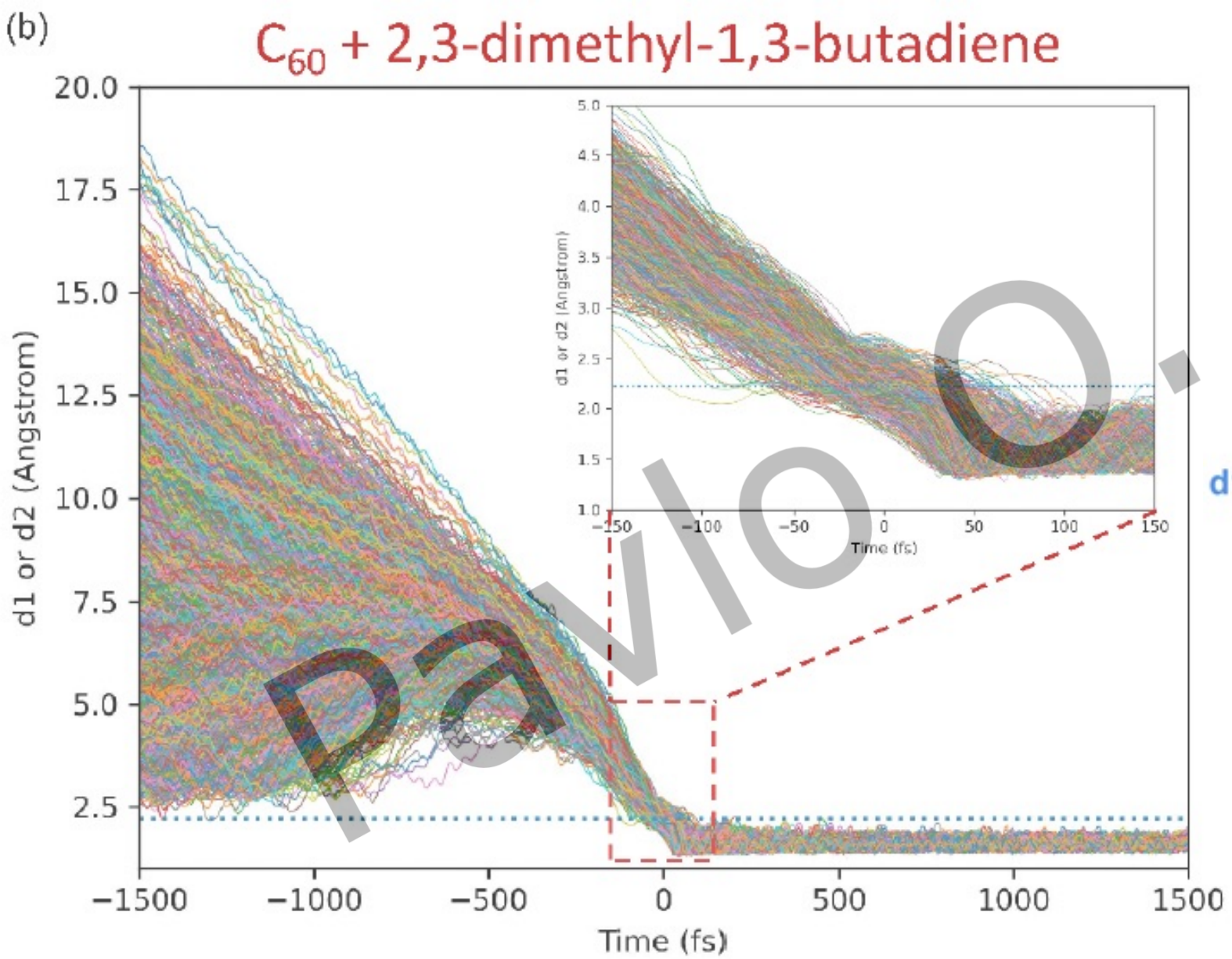


Pavlo O Dral

Y.-F. Hou, Q. Zhang, P. O. Dral. https://doi.org/10.26434/chemrxiv-2024-hwsl4.


(a)
ethene + 1,3-butadiene
d1 or d2 (Angstrom)
Time (fs)
6
5
4
3
2
1
−150
−100
−50
0
50
100
150

Pavlo O. Dral


(b)
C60 + 2,3-dimethyl-1,3-butadiene
d1 or d2 (Angstrom)
Time (fs)
20.0
17.5
15.0
12.5
10.0
7.5
5.0
2.5
−1500
−1000
−500
0
500
1000
1500
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
−150
−100
−50
0
50
100
150

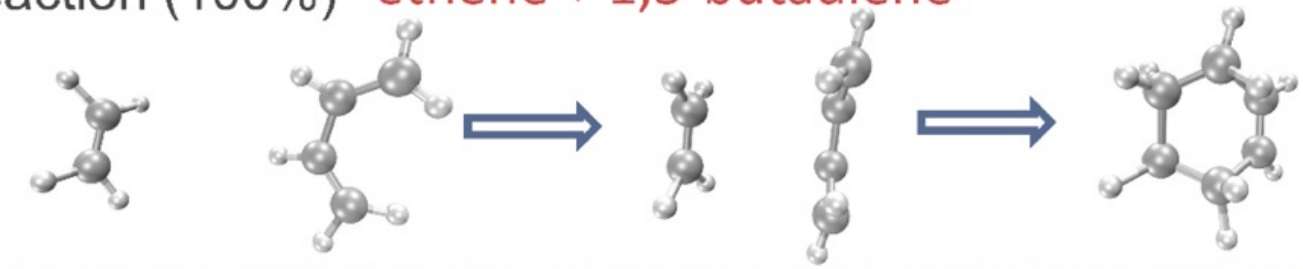

d1
d2
d1 = d2 = 2.227 Å

Pavlo O. Dral

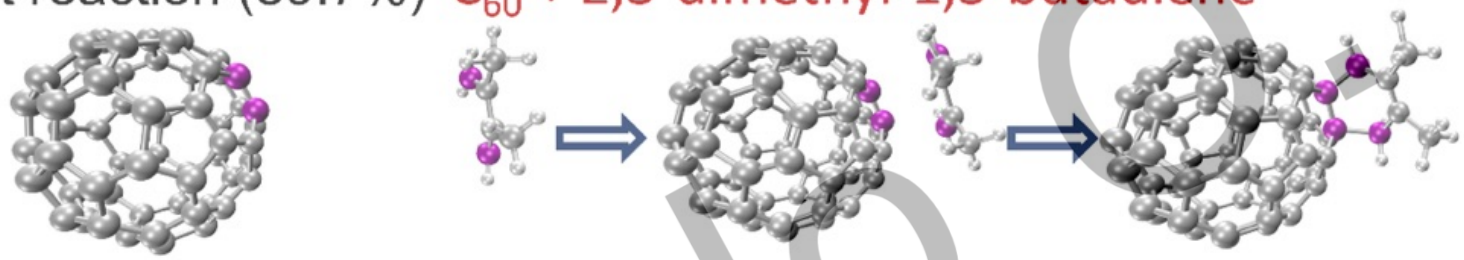
Time-resolved mechanisms

~3000 training points with ANI potential

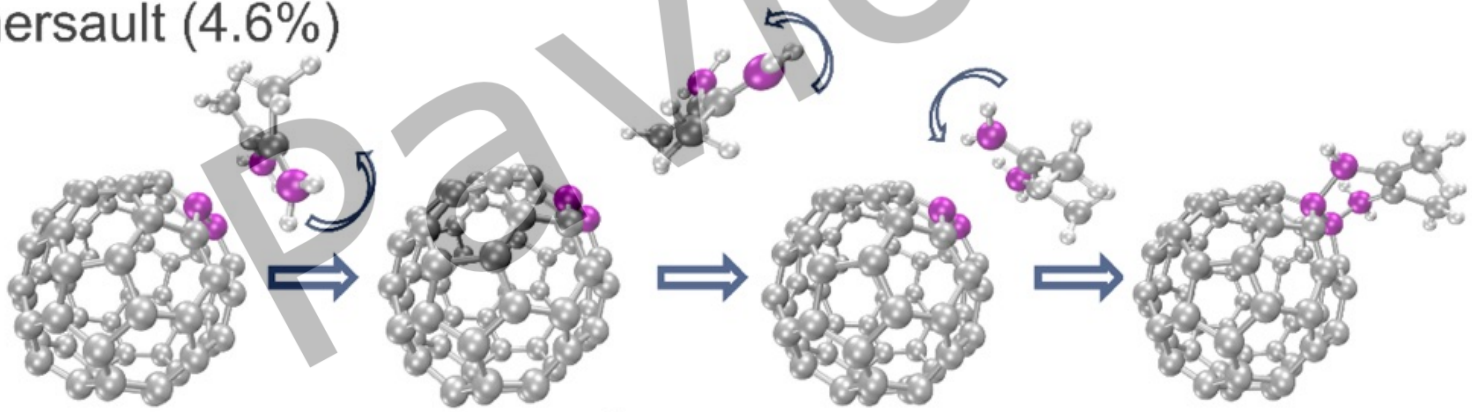
Direct reaction (100%) ethene + 1,3-butadiene



Direct reaction (89.7%) C_{60} + 2,3-dimethyl-1,3-butadiene



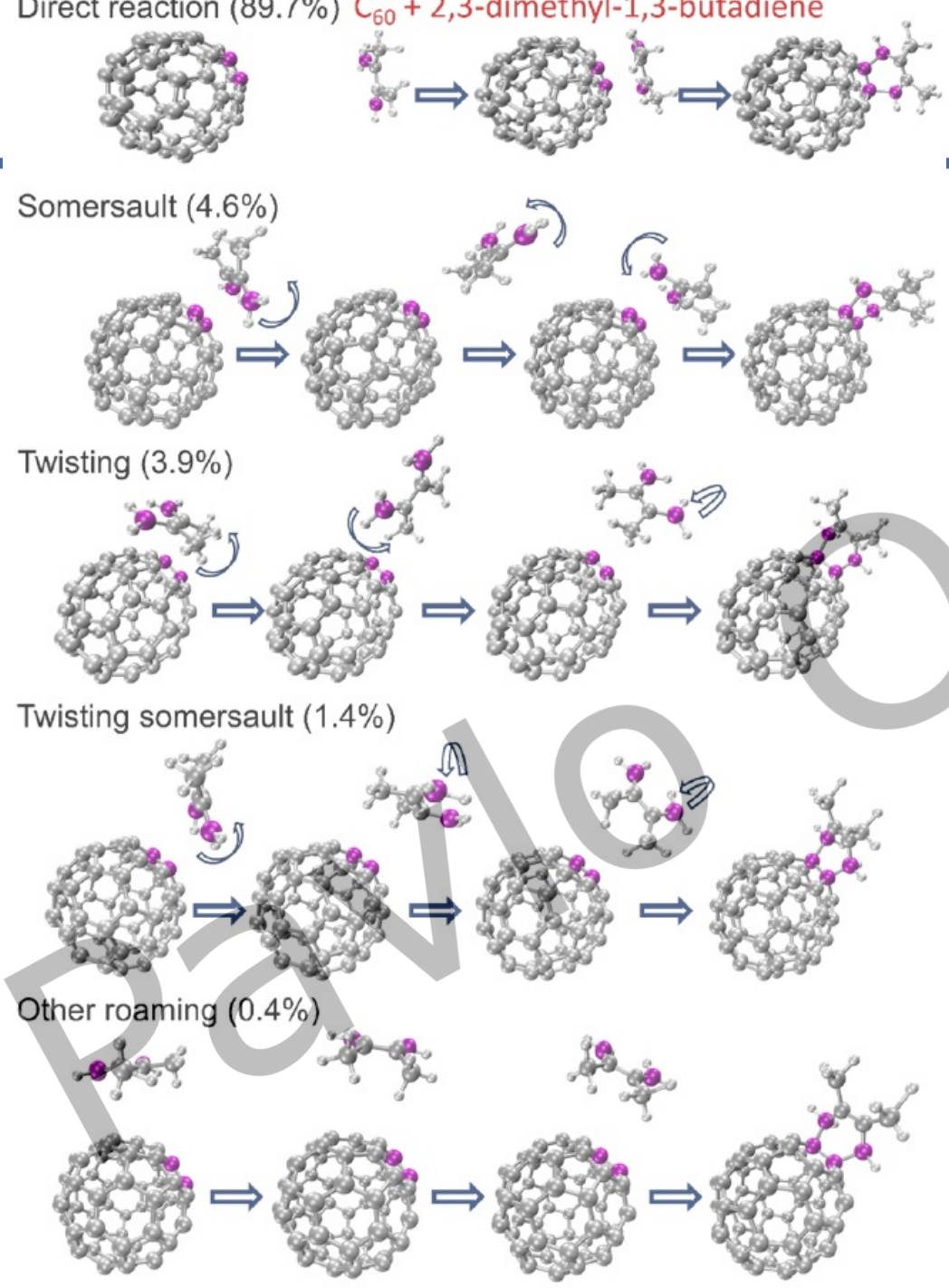
Somersault (4.6%)



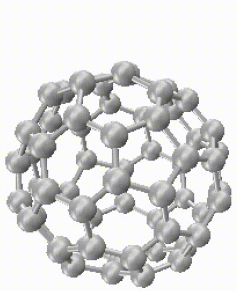
Pavlo O. Dral

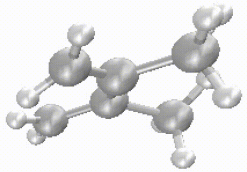
Y.-F. Hou, Q. Zhang, P. O. Dral. https://doi.org/10.26434/chemrxiv-2024-hwsl4.

Time-resolved mechanisms



~3000 training points with ANI potential





Y.-F. Hou, Q. Zhang, P. O. Dral. https://doi.org/10.26434/chemrxiv-2024-hwsl4.

Pavlo O. Dral

Harmonic quantum Boltzmann distribution

- The quantum number n_i of normal mode i is sampled from harmonic quantum Boltzmann distribution:

$$p(n_i) = e^{-n_i h\nu_i / k_B T}\left(1 - e^{-h\nu_i / k_B T}\right)$$

- The coordinates Q_i and momenta P_i are then generated by

$$Q_i = A_i \cos(2\pi R_i),$$
$$P_i = -\omega_i A_i \sin(2\pi R_i)$$

where $\omega_i = 2\pi\nu_i$, $A_i = \sqrt{2E_i}/\omega_i$, $E_i = (n_i + 1/2)h\nu_i$ and R_i is a uniform random number on $[0, 1]$.

- The reaction coordinate (corresponding to the normal mode with the imaginary frequency) is fixed while the mass-weighted momentum P is calculated by $\pm\sqrt{-2k_B T \ln(1 - R)}$, where R is also a uniform random number on $[0, 1]$.

Pavlo O. Dral

Example ML-TS-3.

Run several quasi-classical trajectories for the Diels–Alder reaction starting from the TS.

Analyze where this dynamics leads. Is it consistent with the previous analysis with normal modes and IRC?

```
MD
MLmodelType=ANI
MLmodelIn=da_energy_ani.npz
initConditions=harmonic-quantum-boltzmann
normalModeFile=da_eqmol.json
initTemperature=300
trun=150
dt=0.5
```

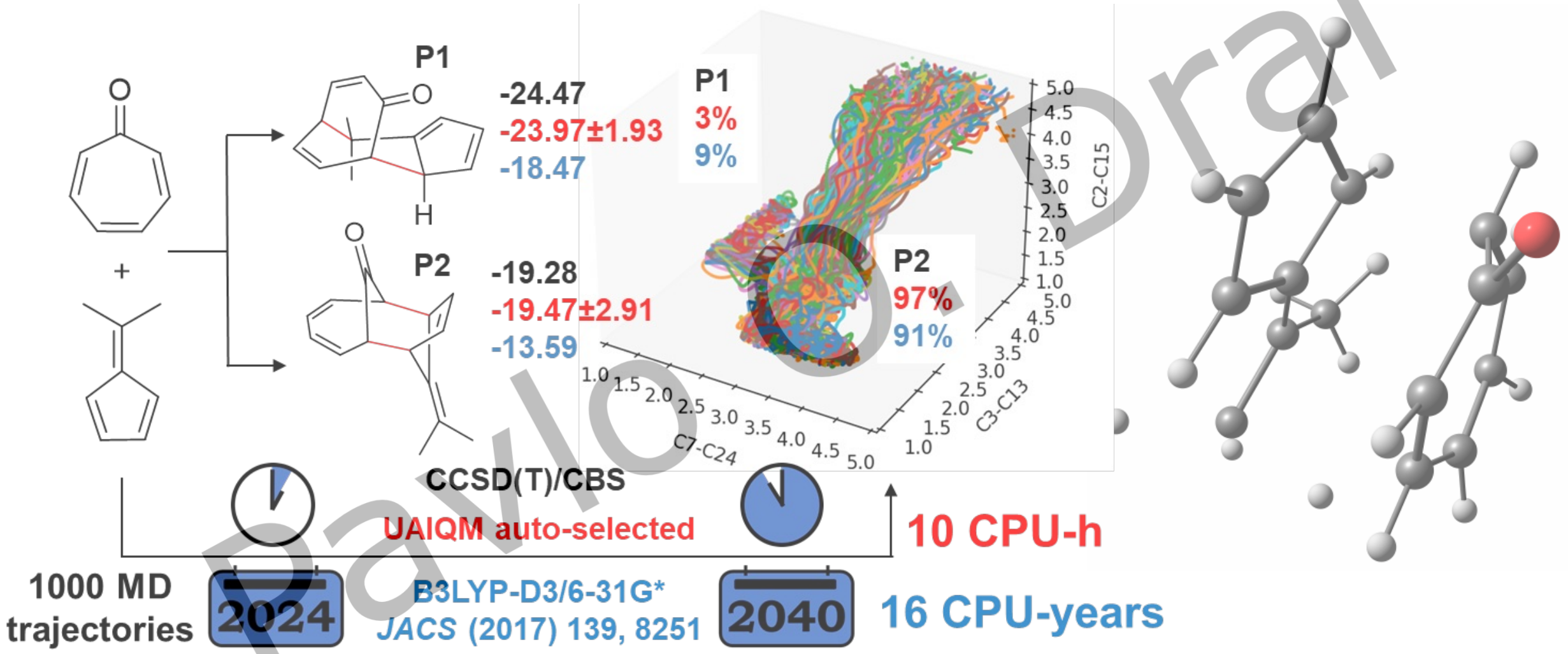
For this dynamics, you need:

- da_eqmol.json - file with frequencies and normal modes: (you can generate such files when you perform `freq` calculations with MLatom, they will be stored in `freq1.json`).
- da_energy_ani.npz - ANI model pretrained on B3LYP/6-31G* data for you. You can use your own level of theory but this dynamics might be quite slow while ML model is much faster.

If you use the same initial conditions but flip the sign in velocities, you can get the set of forward and backward trajectories which you can later put together in one trajectory. If this trajectories goes from reactants to products (or vice versae), it is reactive but some small fraction of trajectories will be non-reactive (i.e., starting and ending with reactants or products).

Pavlo O. Dral

UAIQM: Universal and Updatable AI-QM models



Pavlo O. Dral

Y. Chen, Y.-F. Hou, O. Isayev, P. O. Dral. *Universal and Updatable Artificial Intelligence-Enhanced Quantum Chemical Foundational Models*. **2024**, *submitted*. https://doi.org/10.26434/chemrxiv-2024-604wb.

8. Charge-shift bonding in propellane

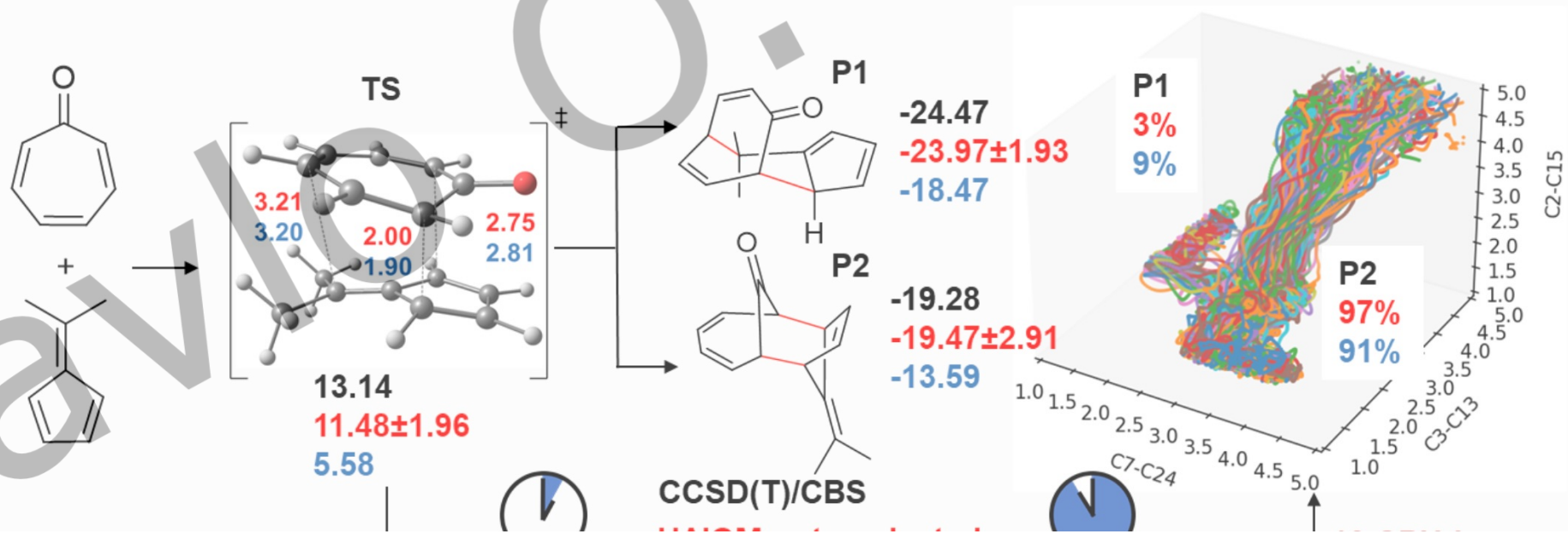
MACHINE LEARNING

1. Machine learning basics
2. ML for PES
3. Universal machine learning models
4. Spectroscopy
5. Transfer learning
6. Molecular dynamics
7. Transition state search and analysis
   - 7.1. TS optimization
   - 7.2. TS validation with normal modes
   - 7.3. Intrinsic reaction coordinate (IRC)
   - 7.4. TS analysis via downhill dynamics
   - 7.5. Quasiclassical trajectories analysis on the bifurcating pericyclic reaction

7.5. Quasiclassical trajectories analysis on the bifurcating pericyclic reaction

If you ever wondered how to get the computation of the beautiful reaction mechanisms, below are instructions from MLatom's documentation prepared by Yuxinxin Chen.

In this section, we will reproduce the quasiclassical trajectories for bifurcating pericyclic reaction reported in JACS (2017) 139, 8251, where the human experts manually chosen the B3LYP-D3/6-31G* level of theory and got the distribution of the products with the resulting 117 reactive trajectories. With the UAIQM models, we can quickly shoot 1000 trajectories and get 867 reactive ones, while the energy profile is more accurate than B3LYP which is known for low reliability on energy prediction.



Pavlo O. Dral

```python
import mlatom as ml

# load initial transitiona state
init_ts1 = ml.data.molecule.from_xyz_file('TS1.xyz')

# choose optimal UAIQM method
uaiqm_optimal = ml.models.uaiqm(method='uaiqm_optimal', verbose=False)
uaiqm_optimal.select_optimal(
    molecule=init_ts1,
    nCPUs=1,
    time_budget='0.1s'
)

# optimize geometry and get frequencies
geomopt = ml.optimize_geometry(
    model=uaiqm_optimal,
    initial_molecule=init_ts1,
    ts=True,
    program='geometric'
)
optmol_ts1 = geomopt.optimized_molecule
ml.freq(model=uaiqm_optimal, molecule=optmol_ts1)

# get initial conditions
for ifreq in range(len(optmol_ts1.frequencies)):
    if optmol_ts1.frequencies[ifreq] > 0 and optmol_ts1.frequencies[ifreq] < 100:
        optmol_ts1.frequencies[ifreq] = 100
```

... Long script
Will take you 10 hours to finish...

Pavlo O. Dral