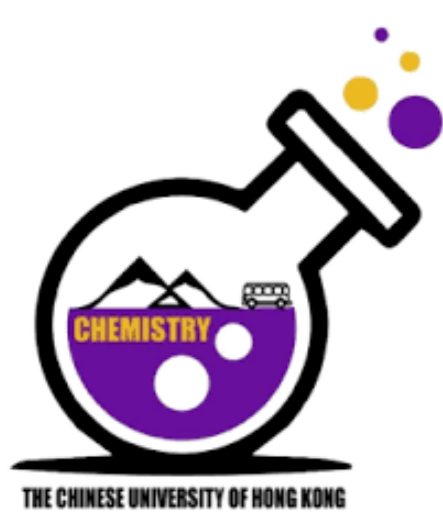


# CHEMSMART: Chemistry Simulation and Modeling Automation Toolkit for High-Efficiency Computational Chemistry Workflows

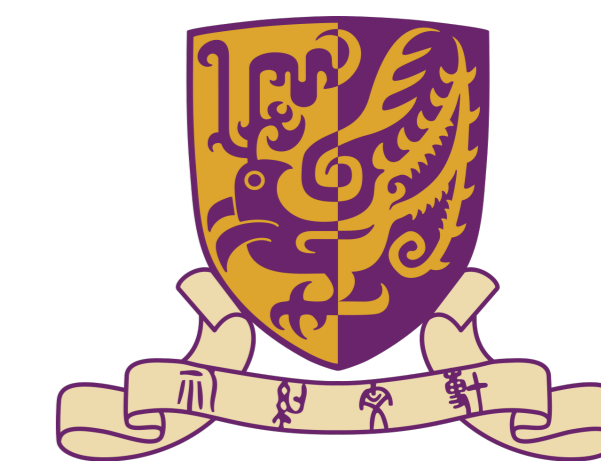
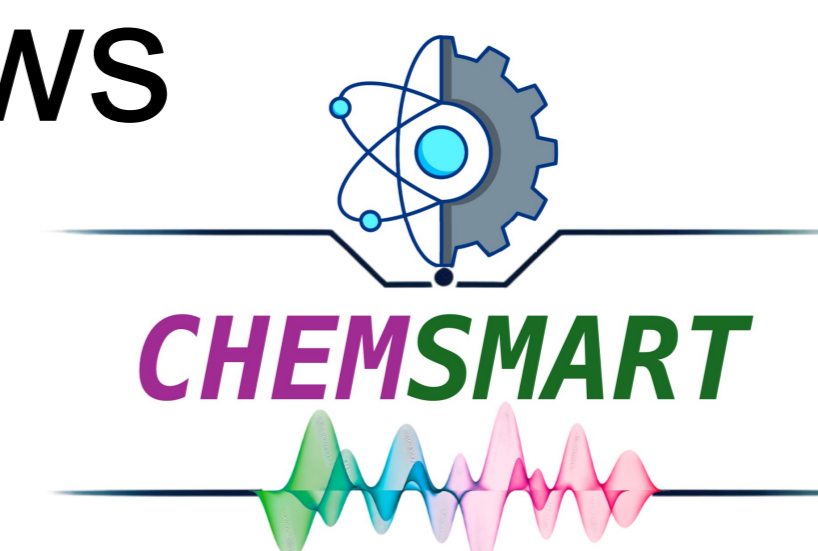


Jingyi Liu<sup>1</sup>, Huiwen Tan<sup>1</sup>, Zihan Li<sup>1</sup>, Lewen Wang<sup>1,2</sup>, Benjamin W. J. Chen<sup>3</sup>, Xinglong Zhang<sup>1,\*</sup>

<sup>1</sup>Department of Chemistry, the Chinese University of Hong Kong. Email: [xinglong.zhang@cuhk.edu.hk](mailto:xinglong.zhang@cuhk.edu.hk)

<sup>2</sup>Department of Chemistry, Hong Kong Baptist University

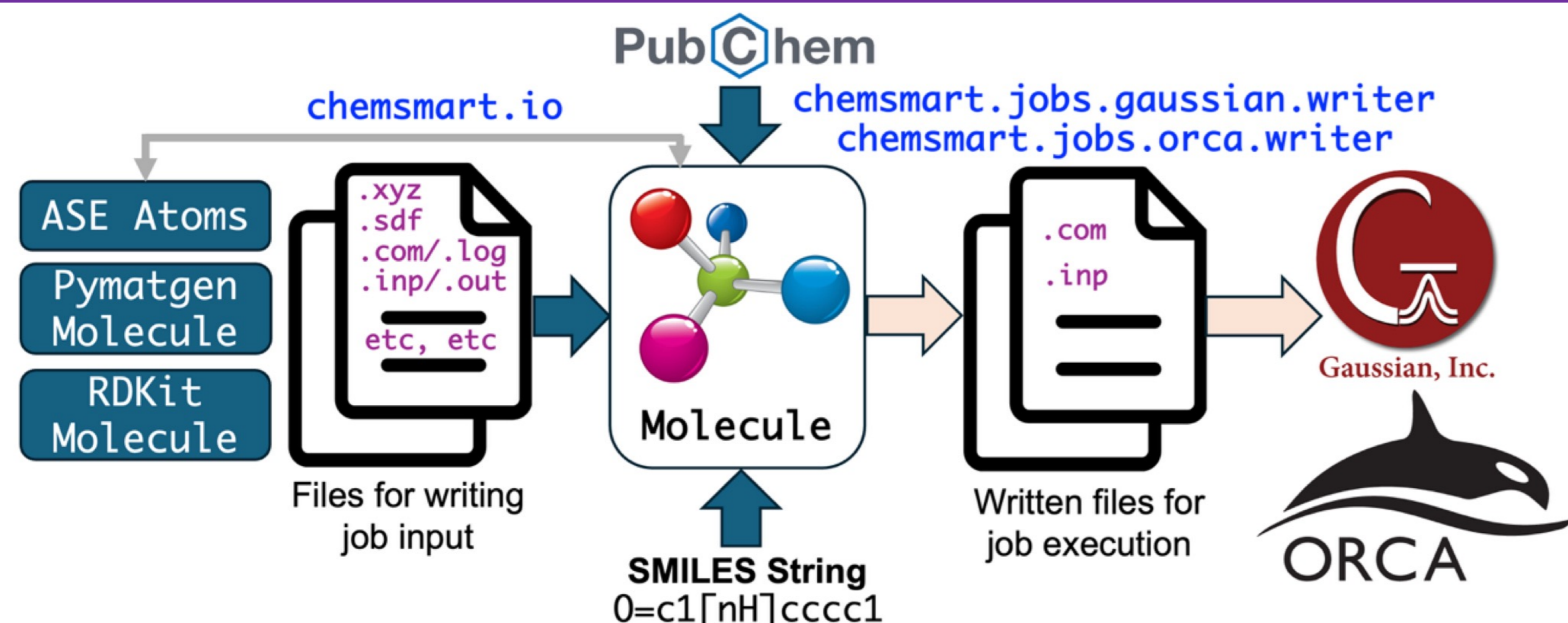
<sup>3</sup>Institute of High Performance Computing, Agency for Science, Technology and Research (A\*STAR)



## Introduction

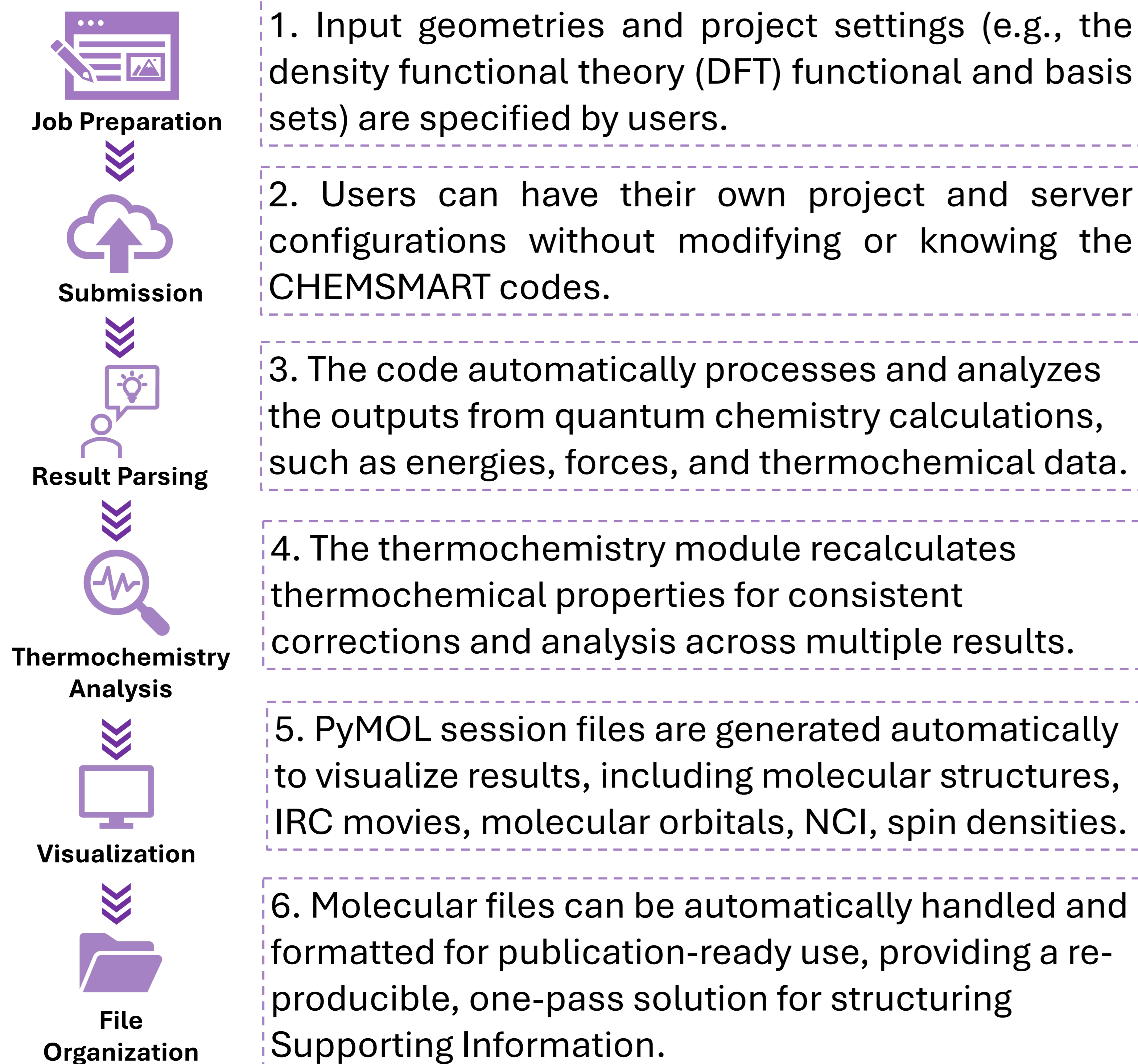
- 1) Computational chemistry tasks—geometry optimization, transition states, thermochemistry, noncovalent analysis—are typically fragmented across multiple codes and manual scripts.
- 2) Compared to that of heterogeneous catalysis, automation frameworks for molecular systems are far less developed.
- 3) CHEMSMART provides a **unified, extensible integration for end-to-end molecular quantum chemistry workflows**, ensuring reproducibility and reducing human intervention.

## Package Design



- 1) **Python-based modular design:** individual modules for structure preparation, job execution, data parsing, and visualization.
- 2) **Central Molecule Object:** encapsulates coordinates, charge, multiplicity, and metadata, serving as the data backbone.
- 3) **Interoperability Layer:** interfaces with common quantum chemistry engines (e.g. Gaussian, ORCA, etc.).
- 4) **CLI and scripting support:** tasks can be executed interactively or via batch workflows.

## Modular Architecture

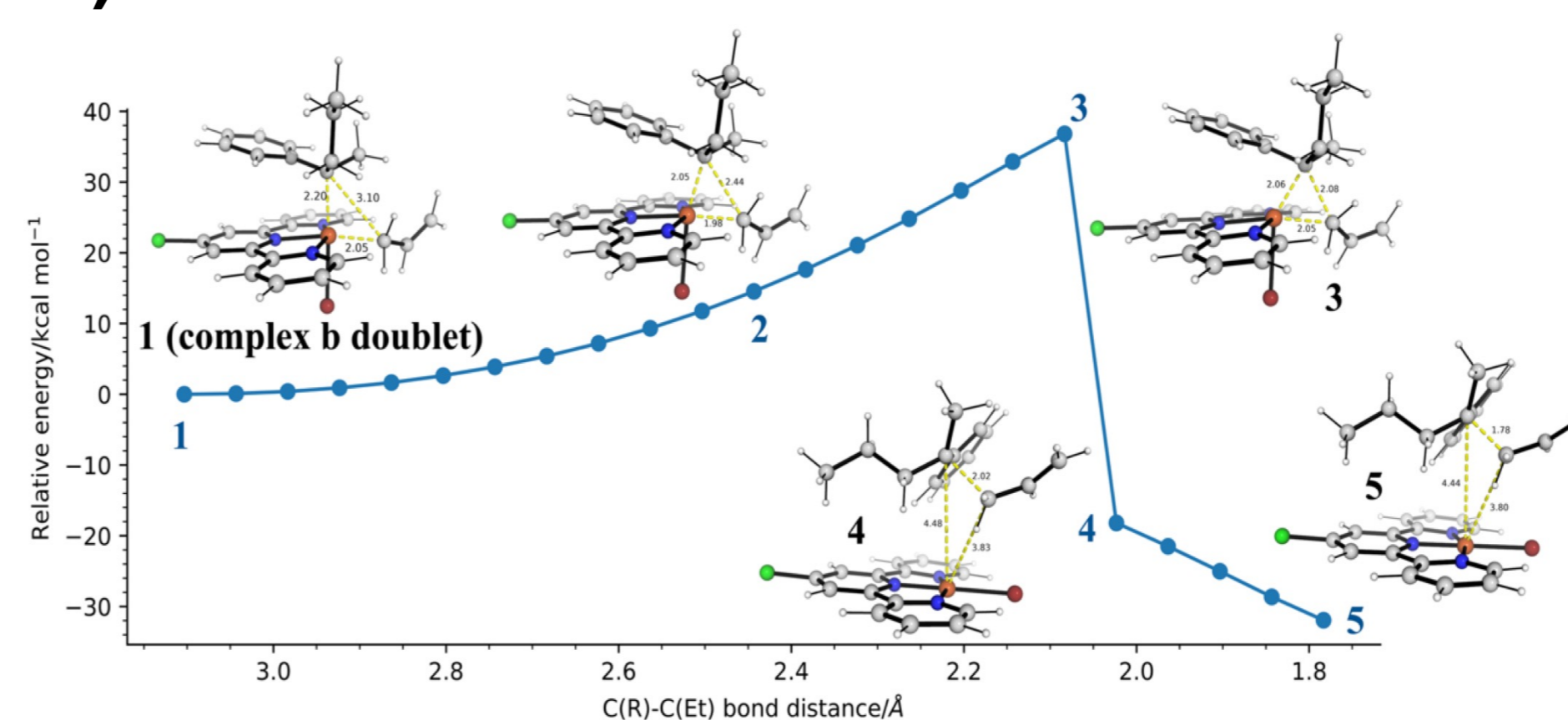


## Future Development

- 1) Integration with additional software.
- 2) Multi-scale simulation (QM/MM) workflows, including integration with machine learning interatomic potentials (MLIPs).
- 3) Data standardization aligned with FAIR data principles (Findable, Accessible, Interoperable, Reusable).

## Use Cases

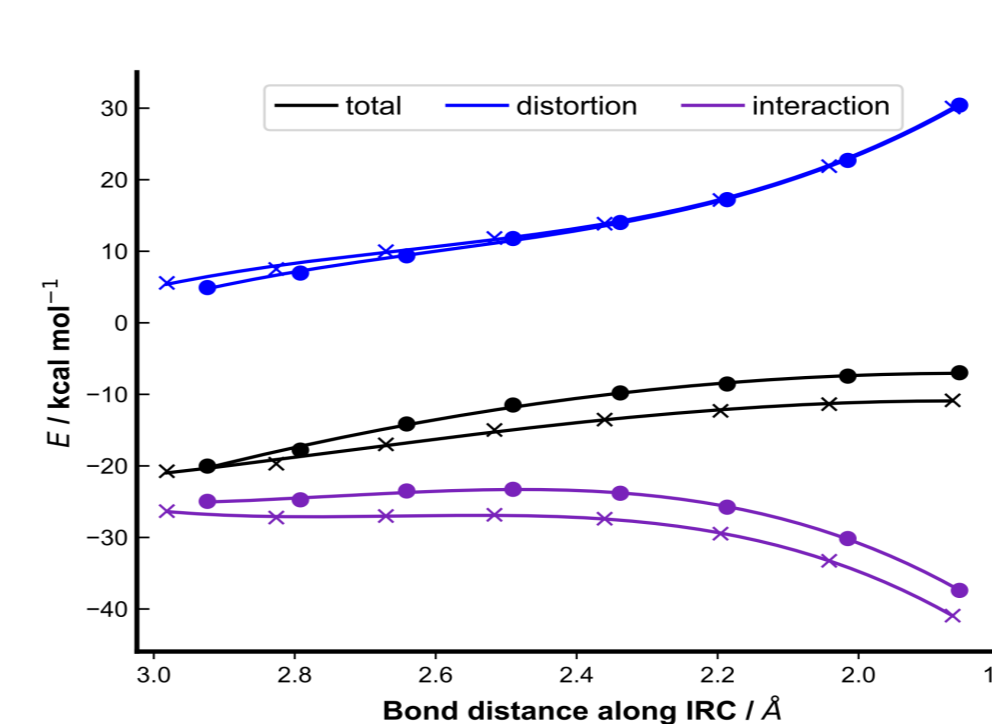
### 1) PES scan



The scanned coordinate (via `-c` flag) here is bond distance between atoms numbered 1 and 2, with a step size of  $-0.1 \text{ \AA}$  (via stepsize `-s` flag) with a total of 20 steps (via num-steps `-n` flag).

```
chemsmart sub gaussian -p iron -f structure.com scan -c [1,2] -s -0.1 -n 20
```

### 2) Distortion-interaction / Activation strain analysis



The code allows users to perform DI/AS analyses along the IRC paths of two competing transition states and to calculate the distortion and interaction at each point. Here `-i` specifies the number of atoms belong to one component of the TS.

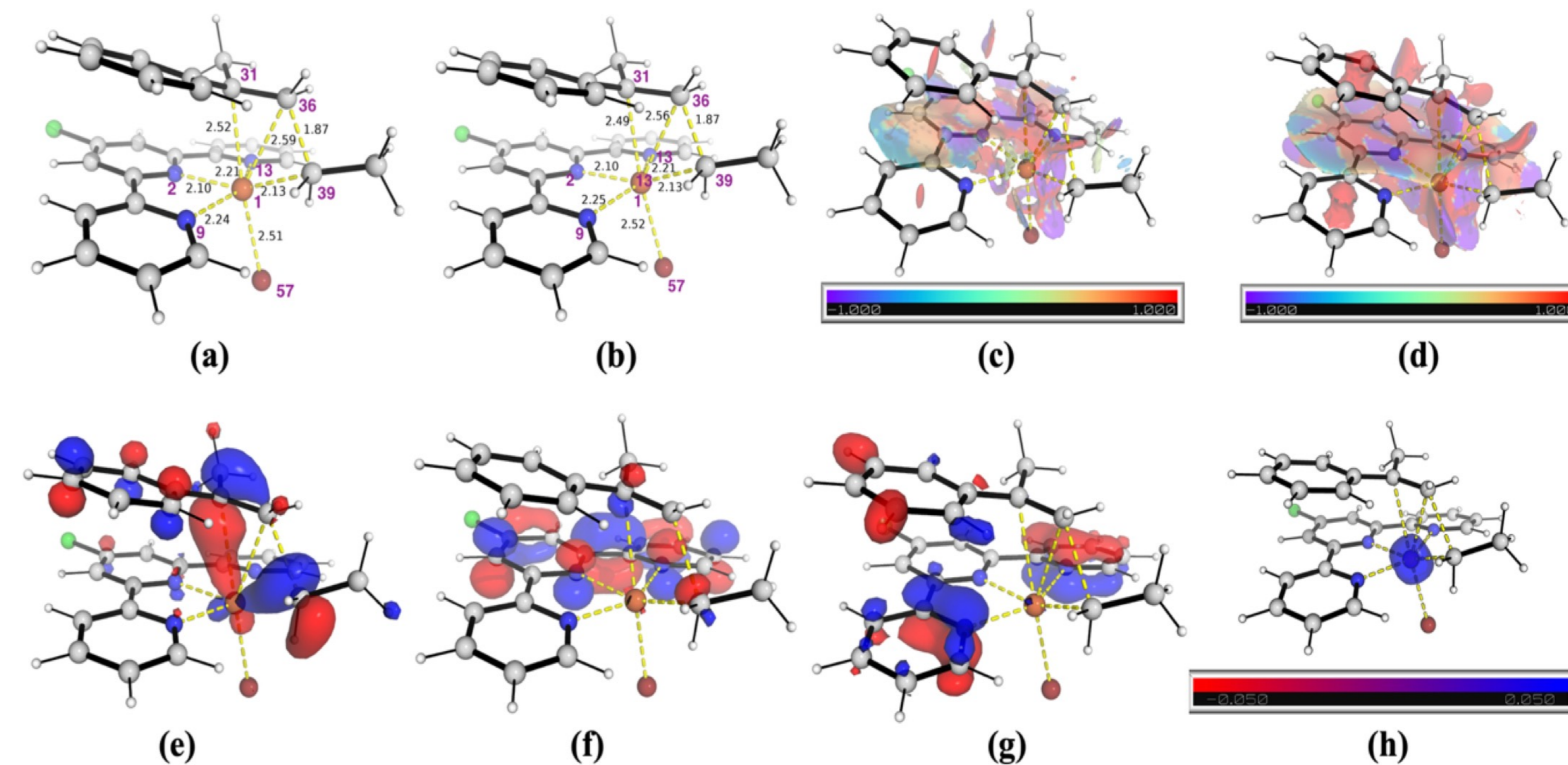
```
chemsmart sub gaussian -p iron -f ircf.log dias -i 1-20,24,30  
plot_dias.py -p gaussian -a 5 -b 7
```

### 3) Thermochemical calculation with quasi-RRHO corrections

```
chemsmart run thermochemistry -T 298.15 -f water_mp2.log -csg 100  
chemsmart run thermochemistry -T 298.15 -f water_mp2.log -ch 200  
chemsmart run thermochemistry -T 298.15 -f water_mp2.log -cst 40 -ch 100
```

The thermochemistry command enables users to recalculate and apply corrections to thermochemical properties. The `-csg/-ch/-cst` options apply quasi-RRHO corrections with a specified frequency cut-off (in  $\text{cm}^{-1}$ ) to calculate the corrected entropy (qh-S), enthalpy (qh-H), and Gibbs free energy (qh-G(T)).

### 4) Visualization and molecular analysis



Transition states geometries can be visualized using the `mol` command, with `-c` specifying the bonds to be displayed (a, b). Non-covalent interaction (NCI) commands generate NCI plots from `.wfn` files (c) and from promolecular density (d). Molecular orbitals can be plotted using the `mo` command (e, f, g), while spin density distribution can be visualized with the `spin` command (h).

```
chemsmart run mol -f iron_quintet_modred_ts.log visualize -c  
[[1,2],[1,9],[1,13],[1,31],[1,36],[1,39],[1,57],[36,39]]  
chemsmart sub gaussian -f iron_quintet_modred_ts.log nci  
chemsmart sub nciplot -f ts_nci.wfn  
chemsmart run mol -f iron_quintet_modred_ts.log nci  
chemsmart run mol -f iron_quintet_modred_ts.log mo --homo/--lumo  
chemsmart run mol -f iron_quintet_modred_ts.log spin
```

## References:

(1) Zhang, X.; Tan, H.; Liu, J.; Li, Z.; Wang, L.; Chen, B. W. J. CHEMSMART: Chemistry Simulation and Modeling Automation Toolkit for High-Efficiency Computational Chemistry Workflows. *arXiv* 2025, arXiv:2508.20042.



Github Repo Documentation