NIBR

Building quantum chemistry pipelines with RDKit

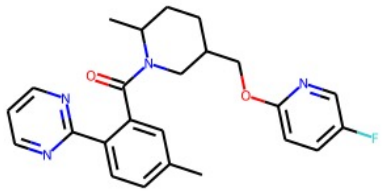
Oct 2021

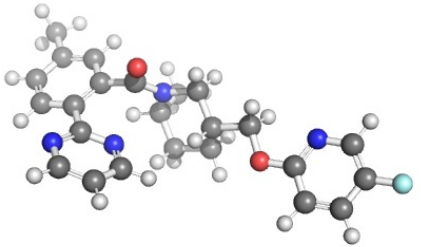
Jimmy Kromann, Novartis

NOVARTIS | Reimagining Medicine

What do I want to show

I constantly need to go between Quantum Chemistry and Cheminfomatics





And I was writing the same functions again and again (hacky .sh and .py files)

I need

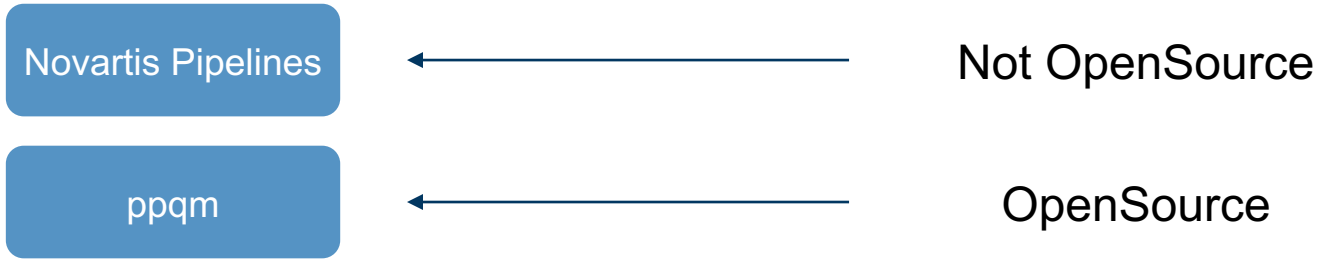
- To programmatically need administrate many compounds (and their conformers)
- To be able to run QM calculations
- To read properties from QM calculations
- Retain cheminfomatic information
- So that I can make decisions on the properties
  - (data science, drug discovery, org-chemistry, etc)

Implementation idea

- Do **not** create a yet another molecule class
  - Extend functionality by `functionality(rdkit_molobj)`
- QM Expert level (and easy) interface
- Sane defaults for non-experts

- Python module for pipeline
  - Stable
  - Easy to manage
  - Easy to scale
- Jupyter friendly
  - Easy viewer
  - Experimental
  - Interactive

github.com/ppqm/ppqm

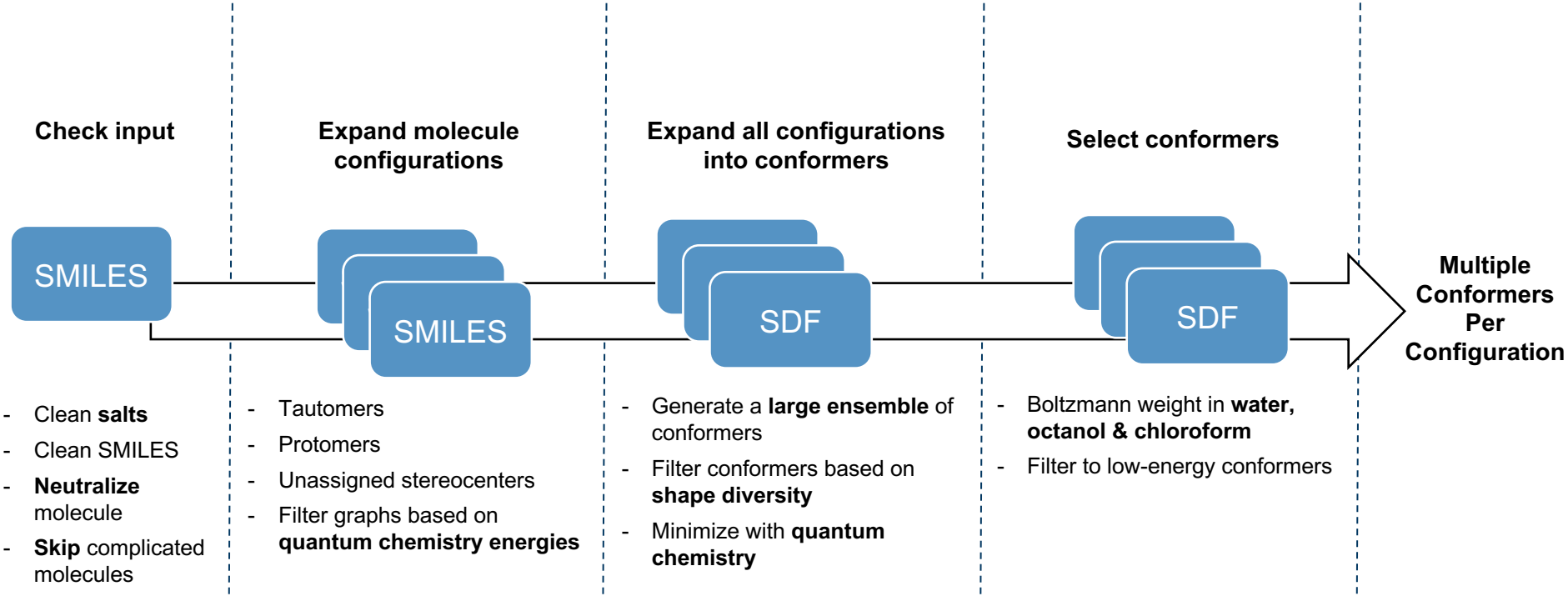


ppqm – Psi Phi Quantum Mechanics

And what do you use it for?

- Productionized molecule pipeline **conformers** and **quantum properties**
- Quantum chemistry **descriptors** for machine learning (xTB properties)

A pipeline for (all) our molecules



Show example

Other xTB and RDKit examples

- Regioselecitivty of aromatic carbons
- Tautomer selections
- Torsional scan
- Fast conformer selection
- …

- Also examples with Gaussian, MOPAC, MNDO and GAMESS

The future of computational chemistry

- No more torsional scans in Gaussview, submit bash scripts and read into Excel

- Future chemistry classes will be all notebooks
  - rdkit
  - ppqm (xTB)
  - sklearn
  - Etc

- Small steps towards

TODO

- mypy
- github actions (only for xtb tasks for now)
- conda-forge package

- Many other QM wrappers out there, maybe need to merge
- Still technical, needs **better unit managment**

Acknowledgements

Novartis

- **Richard Lewis**
- **Rainer Wilcken**
- **Aniko Udvarhelyi**
- **Hagen Muenkler**
- **Riccardo Vianello**
- **Paolo Tosco**
- And **Nik Stiefl**

Open Source community

- **Mads Koerstz** (Uni. Copenhagen)
- **Puck van Gerwen** (EPFL)
- **Anders Christensen**
- **Lars Bratholm**
- **Prof. Casper Steinmann** (Aalborg Uni.)

Mange Tak
Thanks
Danke

Package source code and all notebooks

`github.com/ppqm/ppqm`

:wq

NOVARTIS | Reimagining Medicine