open forcefield

@openforcefield

www.openforcefield.org

BespokeFit

Automated bespoke parameters for OpenFF

Joshua Horton

BespokeFit

Why?

Force Fields

Allow us to parametrise vast amounts of chemical space quickly via:

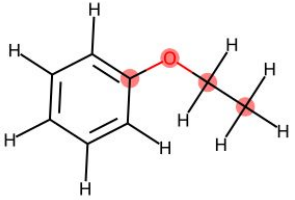
- a large number of atom type specific parameters eg OPLS or GAFF
- or smart typing a smaller collection of parameters using chemical perception (OpenFF)

Force Field	Number of unique torsion parameters
OpenFF-1.0.0	157
OpenFF-1.2.0	163
OpenFF-1.3.0	167
OPLS3	48142
OPLS3e	146,669

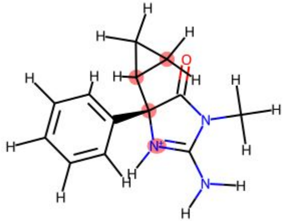
J. Chem. Theory Comput. 2019, 15, 1863−1874 Article

BespokeFit

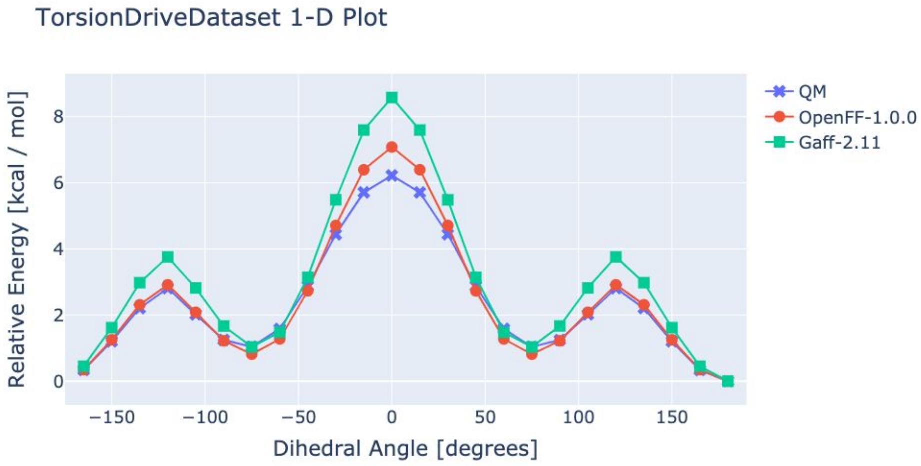
Torsion Parameters

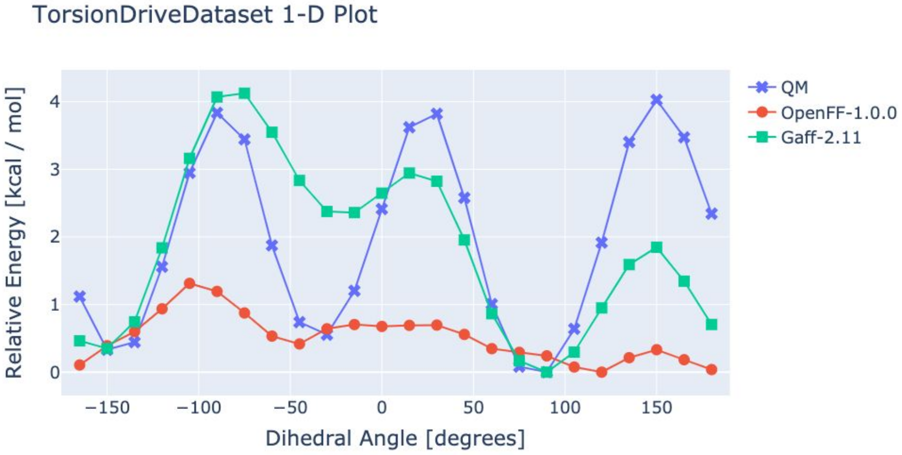


Openff-1.0.0 works well with no modifications on simple molecules.



However it can struggle in some more complex cases seen in the Schrödinger benchmark systems. (BACE)

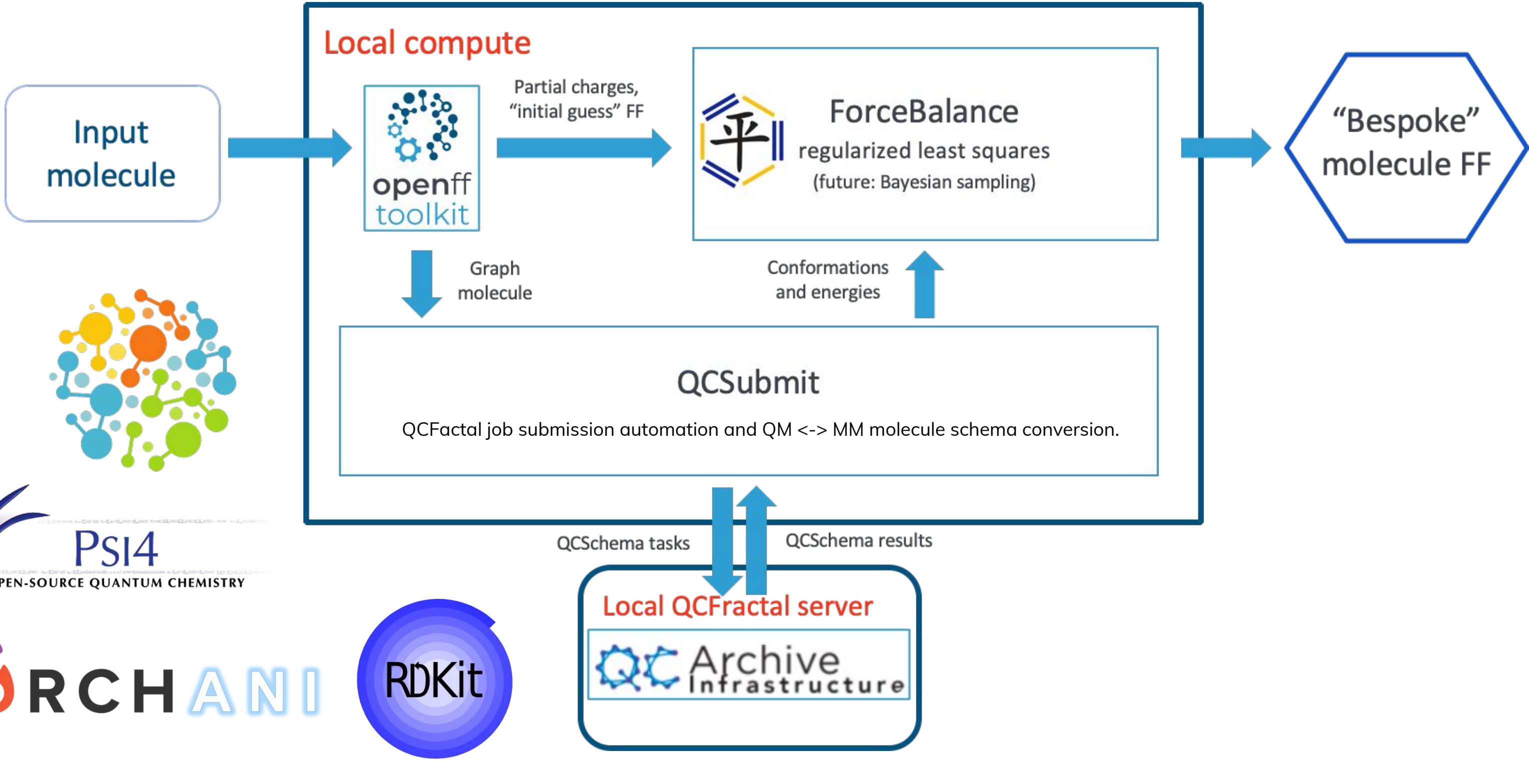

TorsionDriveDataset 1-D Plot


TorsionDriveDataset 1-D Plot

BespokeFit

Bespoke framework

A typical bespoke pipeline powered with QCSubmit allowing for a direct interface with the QC ecosystem. One interface for many programs



Input molecule

Local compute

openff toolkit

Partial charges, “initial guess” FF

ForceBalance

regularized least squares

(future: Bayesian sampling)

“Bespoke” molecule FF

Graph molecule

Conformations and energies

QCSubmit

QCFactal job submission automation and QM <-> MM molecule schema conversion.

QCSchema tasks

QCSchema results

Local QCFractal server

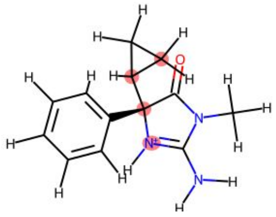
QC Archive Infrastructure

PSI4 OPEN-SOURCE QUANTUM CHEMISTRY

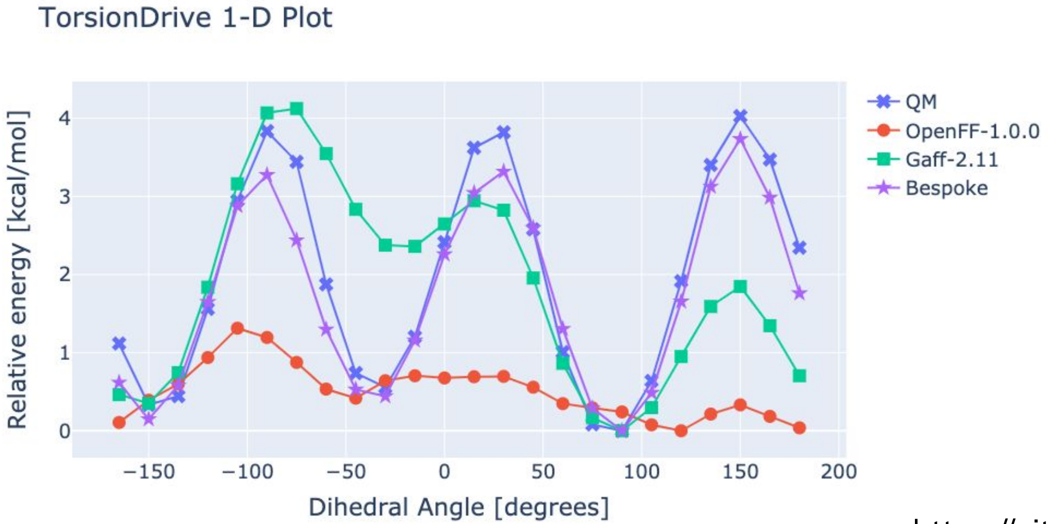
xTB TORCHANI RDKit

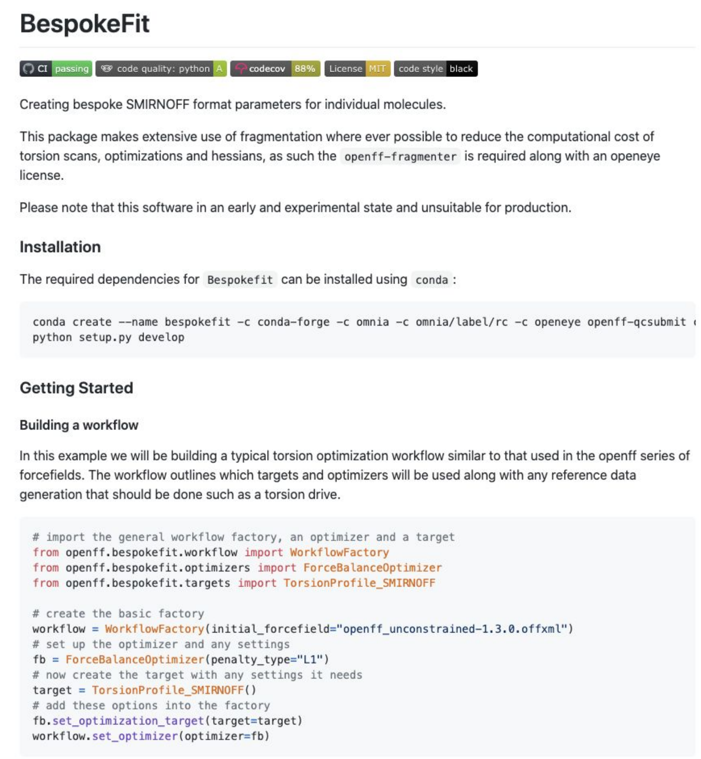
BespokeFit

Result



Clear improvement of the PES after fitting bespoke parameters.





BespokeFit

CI passing | code quality: python A | codecov 88% | License MIT | code style black

Creating bespoke SMIRNOFF format parameters for individual molecules.

This package makes extensive use of fragmentation where ever possible to reduce the computational cost of torsion scans, optimizations and hessians, as such the `openff-fragmenter` is required along with an openeye license.

Please note that this software in an early and experimental state and unsuitable for production.

Installation

The required dependencies for `Bespokefit` can be installed using `conda` :

```
conda create --name bespokefit -c conda-forge -c omnia -c omnia/label/rc -c openeye openff-qcsubmit
python setup.py develop
```

Getting Started

Building a workflow

In this example we will be building a typical torsion optimization workflow similar to that used in the openff series of forcefields. The workflow outlines which targets and optimizers will be used along with any reference data generation that should be done such as a torsion drive.

```
# import the general workflow factory, an optimizer and a target
from openff.bespokefit.workflow import WorkflowFactory
from openff.bespokefit.optimizers import ForceBalanceOptimizer
from openff.bespokefit.targets import TorsionProfile_SMIRNOFF

# create the basic factory
workflow = WorkflowFactory(initial_forcefield="openff_unconstrained-1.3.0.offxml")
# set up the optimizer and any settings
fb = ForceBalanceOptimizer(penalty_type="L1")
# now create the target with any settings it needs
target = TorsionProfile_SMIRNOFF()
# add these options into the factory
fb.set_optimization_target(target=target)
workflow.set_optimizer(optimizer=fb)
```

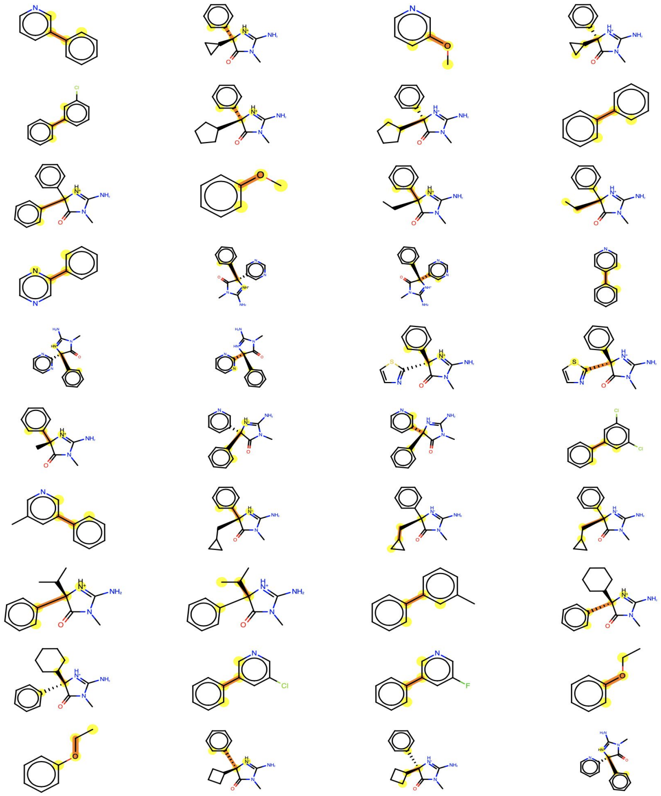
https://github.com/openforcefield/bespoke-fit

BespokeFit

BACE Fragments

Fitting results for the full set of fragments from the BACE Schrödinger set.

Force Field	Average RMSE (kcal/mol)	Average RMSD (angstrom)
OpenFF-1.3.0	0.97	0.17
OpenFF-1.3.0 + Bespoke	0.18	0.13



BespokeFit

Future features

- Bonds
- Angles
- Improper torsions?
- Charges?
- Predicting when to refit parameters?

Beyond

In my new role at Newcastle University I will be continuing my research on the use of bespoke vdW parameters, close range repulsion, virtual site parameters and new functional forms using QUBEKit https://github.com/qubekit/QUBEKit

Validation

- Free energy calculations starting with the TYK2 Schrödinger system.
- Expanding the Lim optimization benchmark to include a bespoke-fit force field.
- Any other important benchmarks to show improvement ?

Demonstration

Now a quick demonstration of easy it is to construct a BespokeFit workflow.

Acknowledgement

- Jamshed Anwar
- Jeff Wagner
- David Mobley
- John Chodera
- Daniel Cole
- OpenFF Team

UK Research and Innovation