

# Molecular Dynamics Simulations of Protein-Ligand Complexes

**Daniel Cole**

School of Natural and Environmental Sciences



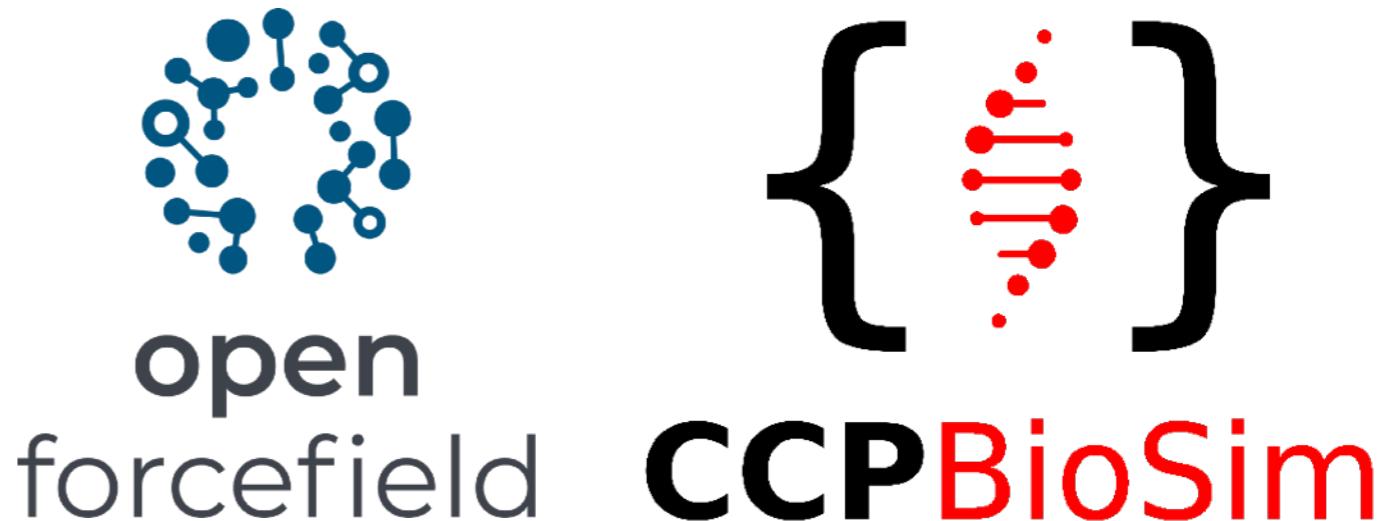
# Acknowledgements

**In person:** Finlay Clark

**Behind the scenes:**

Matt Thompson, Jeff Wagner,  
Lily Wang, Open Force Field

**Organisers:** Charlie Laughton,  
James Gebbie-Rayet,  
Shivani Harshe, Sarah Harris



**UK Research  
and Innovation**

# The protein target

Journal of Medicinal Chemistry > Vol 56/Issue 1 > Article

Subscribed

” Share Jump to Expand

FEATURED ARTICLE | December 17, 2012

## Discovery of Potent Myeloid Cell Leukemia 1 (Mcl-1) Inhibitors Using Fragment-Based Methods and Structure-Based Design

Anders Friberg, Dominico Vigil, Bin Zhao, R. Nathan Daniels, Jason P. Burke, Pedro M. Garcia-Barrantes, DeMarco Camper, Brian A. Chauder, Taekyu Lee, Edward T. Olejniczak, and Stephen W. Fesik\*

View Author Information ▾

Open PDF

SI Supporting Information (1)

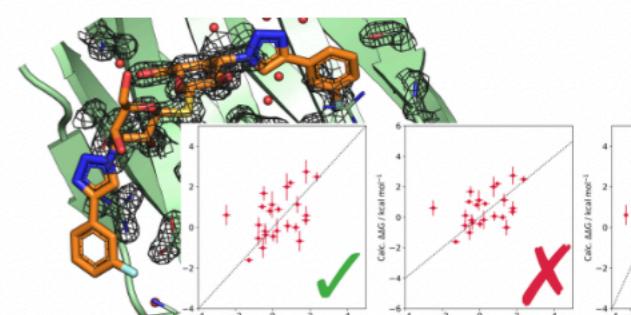
## Best Practices for Constructing, Preparing, and Evaluating Protein-Ligand Binding Affinity Benchmarks [Article v1.0]

David F. Hahn

Computational Chemistry, Janssen Research & Development, Turnhoutseweg 30,  
Beerse B-2340, Belgium

<https://orcid.org/0000-0003-2830-6880>

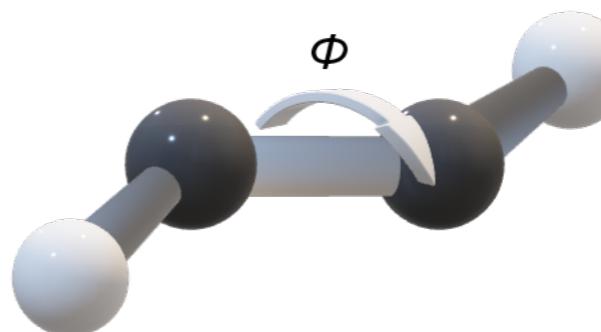
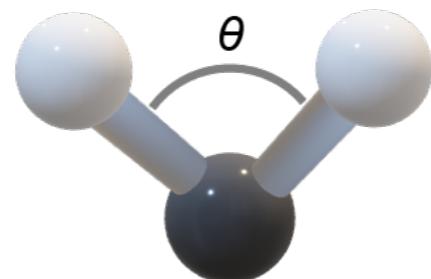
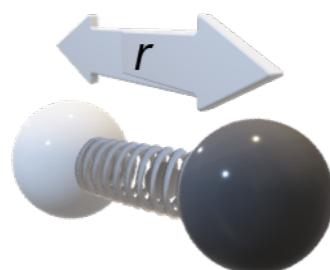
Christopher I. Bayly



# Force Field

Simulations of **large** chemical systems over **long** time scales (e.g. free energy calculations for computer-aided drug design) are crucially reliant on the accuracy of the molecular mechanics force field.

$$E_{Total} = \sum_{Bonds} K_r(r - r_0)^2 + \sum_{Angles} K_\theta(\theta - \theta_0)^2 + \sum_{Torsions} \frac{V_n}{2}[1 + \cos(n\Phi - \gamma)]$$



**Bonded  
(Intramolecular)  
Parameters**

$$+ \sum_{Non-Bonded} \left[ 4\epsilon_{ij} \left\{ \left( \frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left( \frac{\sigma_{ij}}{r_{ij}} \right)^6 \right\} + \frac{q_i q_j}{r_{ij}} \right]$$

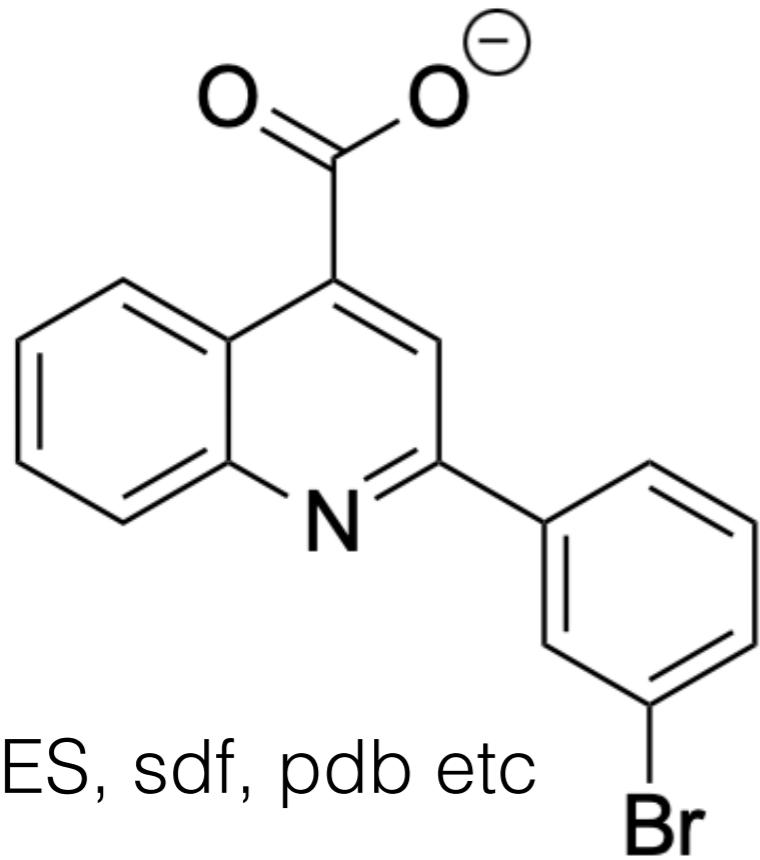


**Non-Bonded  
(Intermolecular)  
Parameters**

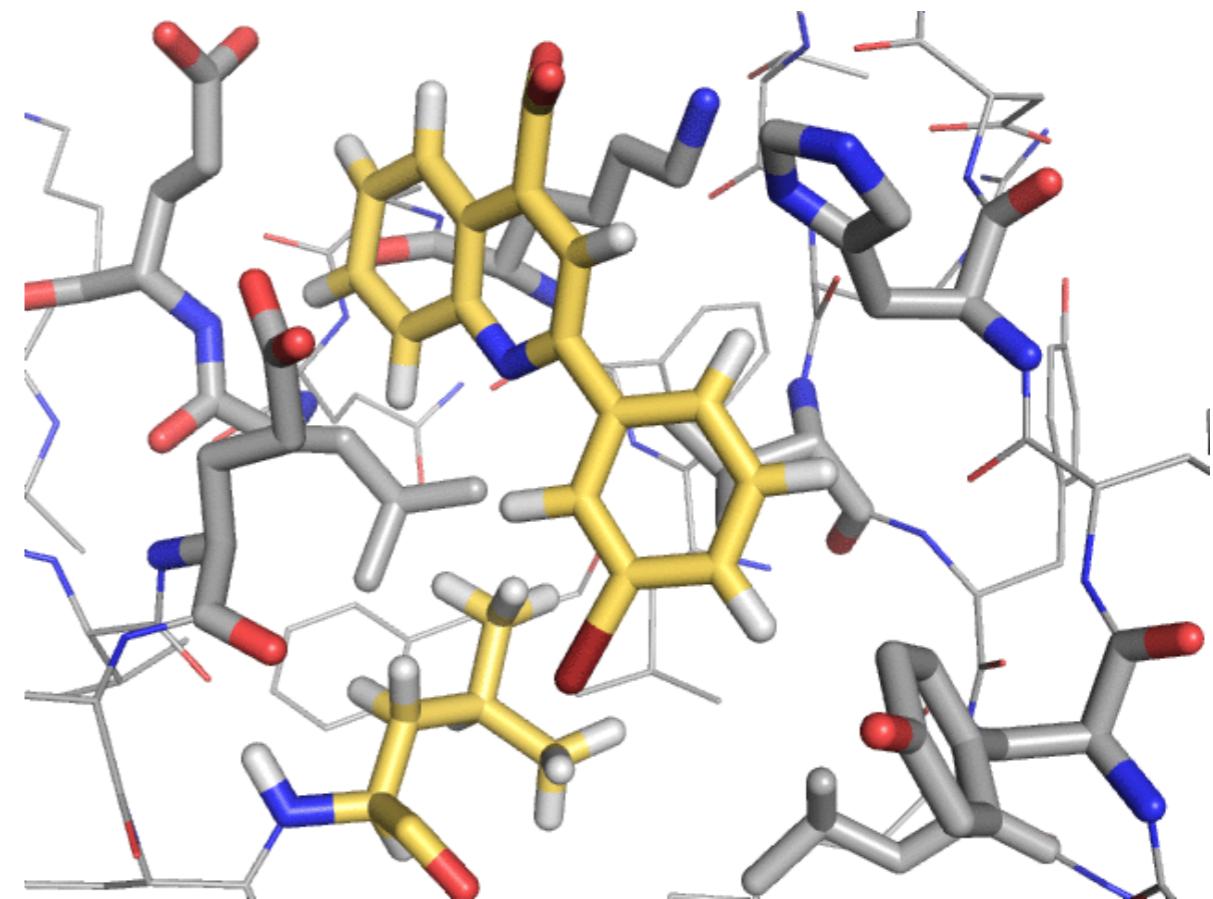
# Force Field

Typically need a protein force field (e.g. AMBER, CHARMM, OPLS), a small molecule force field (e.g. GAFF, CGenFF) and a water force field (e.g. TIP3P).

Can also increasingly use machine learning potentials  
(e.g. ANI, MACE-OFF, Espaloma).

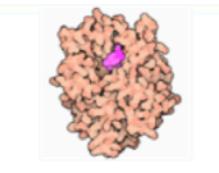


SMILES, sdf, pdb etc



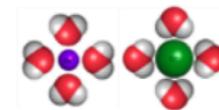
# Force Field

Molecular mechanics force fields have traditionally been products of heroic endeavour (> 100 human-years). How can we bring this into the modern era?

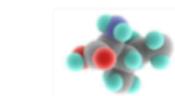


## proteins

post-translational modifications



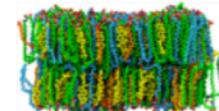
## water ions



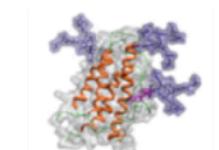
## small molecules



## nucleic acids



## lipids



## carbohydrates

### Amber20 recommendations

J. A. Maier; C. Martinez; K. Kasavajhala; L. Wickstrom; K. E. Hauser; C. Simmerling. ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J. Chem. Theory Comput.*, **2015**, *11*, 3696–3713.

W. D. Cornell; P. Cieplak; C. I. Bayly; I. R. Gould; K. M. Merz, Jr.; D. M. Ferguson; D. C. Spellmeyer; T. Fox; J. W. Caldwell; P. A. Kollman. A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J. Am. Chem. Soc.*, **1995**, *117*, 5179–5197.

N. Homeyer; A. H. C. Horn; H. Lanig; H. Sticht. AMBER force-field parameters for phosphorylated amino acids in different protonation states: phosphoserine, phosphothreonine, phosphotyrosine, and phosphohistidine. *J. Mol. Model.*, **2006**, *12*, 281–289.

H. W. Horn; W. C. Swope; J. W. Pitera; J. D. Madura; T. J. Dick; G. L. Hura; T. Head-Gordon. Development of an improved four-site water model for biomolecular simulations: TIP4P-Ew. *J. Chem. Phys.*, **2004**, *120*, 9665–9678.

I. S. Joung; T. E. Cheatham, III. Molecular dynamics simulations of the dynamic and energetic properties of alkali and halide ions using water-model-specific ion parameters. *J. Phys. Chem. B*, **2009**, *113*, 13279–13290.

P. Li; B. P. Roberts; D. K. Chakravorty; K. M. Merz, Jr. Rational Design of Particle Mesh Ewald Compatible Lennard-Jones Parameters for +2 Metal Cations in Explicit Solvent. *J. Chem. Theory Comput.*, **2013**, *9*, 2733–2748.

J. Wang; R. M. Wolf; J. W. Caldwell; P. A. Kollman; D. A. Case. Development and testing of a general Amber force field. *J. Comput. Chem.*, **2004**, *25*, 1157–1174.

R. Galindo-Murillo; J. C. Robertson; M. Zgarbovic; J. Sponer; M. Otyepka; P. Jureska; T. E. Cheatham. Assessing the Current State of Amber Force Field Modifications for DNA. *J. Chem. Theory Comput.*, **2016**, *12*, 4114–4127.

A. Perez; I. Marchan; D. Svozil; J. Sponer; T. E. Cheatham; C. A. Laughton; M. Orozco. Refinement of the AMBER Force Field for Nucleic Acids: Improving the Description of alpha/gamma Conformers. *Biophys. J.*, **2007**, *92*, 3817–3829.

M. Zgarbova; M. Otyepka; J. Sponer; A. Mladek; P. Banas; T. E. Cheatham; P. Jurecka. Refinement of the Cornell et al. Nucleic Acids Force Field Based on Reference Quantum Chemical Calculations of Glycosidic Torsion Profiles. *J. Chem. Theory Comput.*, **2011**, *7*, 2886–2902.

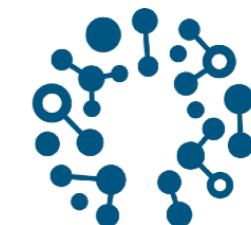
Å. Skjekvik; B. D. Madej; R. C. Walker; K. Teigen. Lipid11: A modular framework for lipid simulations using amber. *J. Phys. Chem. B*, **2012**, *116*, 11124–11136.

C. J. Dickson; B. D. Madej; A. A. Skjekvik; R. M. Betz; K. Teigen; I. R. Gould; R. C. Walker. Lipid14: The Amber Lipid Force Field. *J. Chem. Theory Comput.*, **2014**, *10*, 865–879.

K. N. Kirschner; A. B. Yongye; S. M. Tschampel; J. González-Outeiriño; C. R. Daniels; B. L. Foley; R. J. Woods. GLYCAM06: A generalizable biomolecular force field. Carbohydrates. *J. Comput. Chem.*, **2008**, *29*, 622–655.

# Open Force Field

The Open Force Field Initiative is an international academic-industry collaboration working together to advance science & infrastructure required to build force fields.



open  
forcefield



## OPEN SOFTWARE

Automated infrastructure enables rapid experimentation with minimum human intervention



## OPEN DATA

Access to large, high quality experimental and quantum chemical data facilities easy curation of balanced train / test sets



## OPEN SCIENCE

Exploring new force field science:  
**hypothesis - build software - train - test - iterate**  
is now almost routine



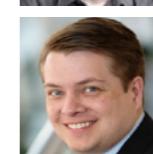
John Chodera (MSKCC)



Michael Gilson (UCSD)



David Mobley (UC Irvine)



Michael Shirts (CU Boulder)



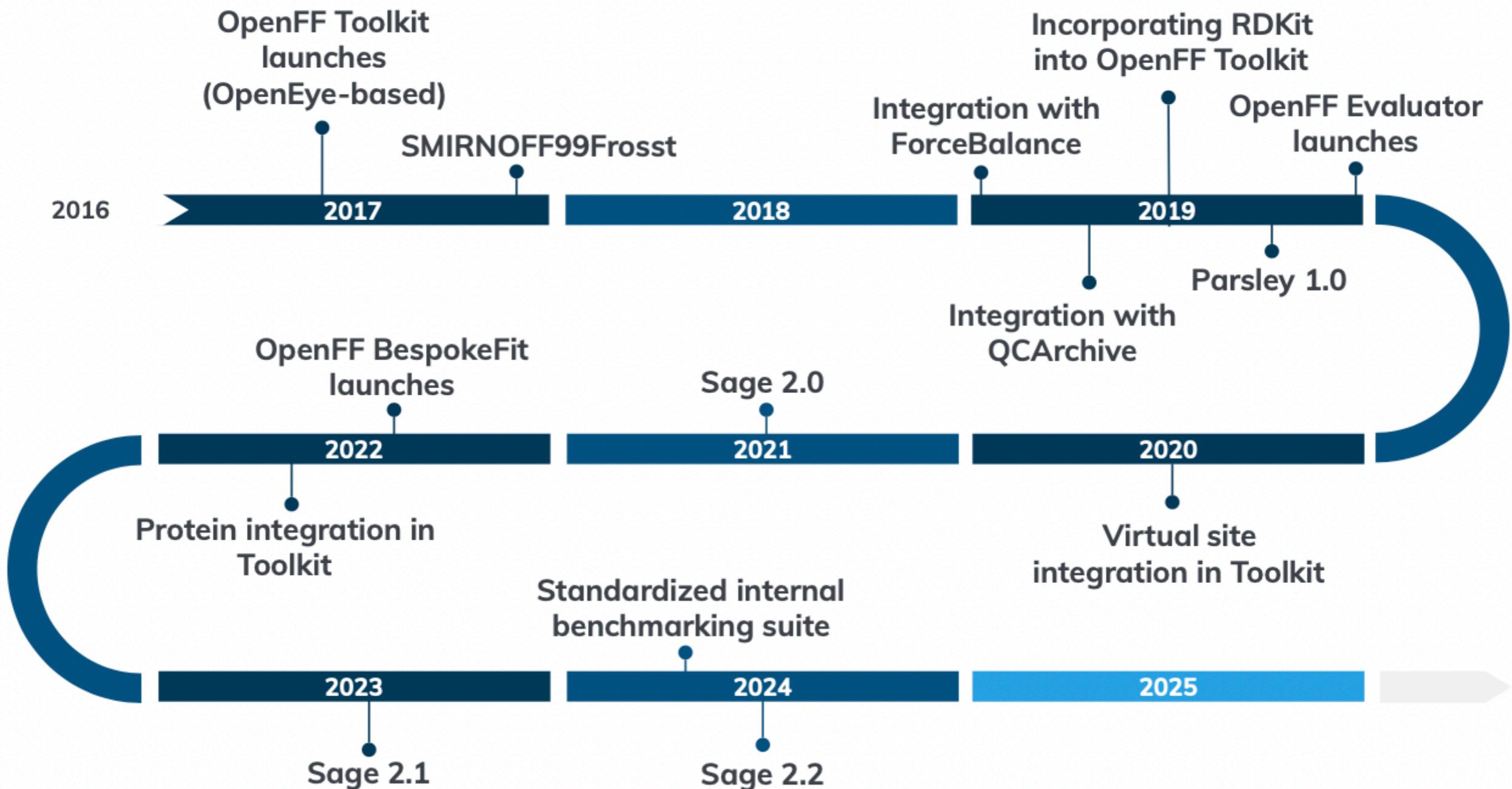
Jeff Wagner  
Technical Lead



Lily Wang  
Science Lead

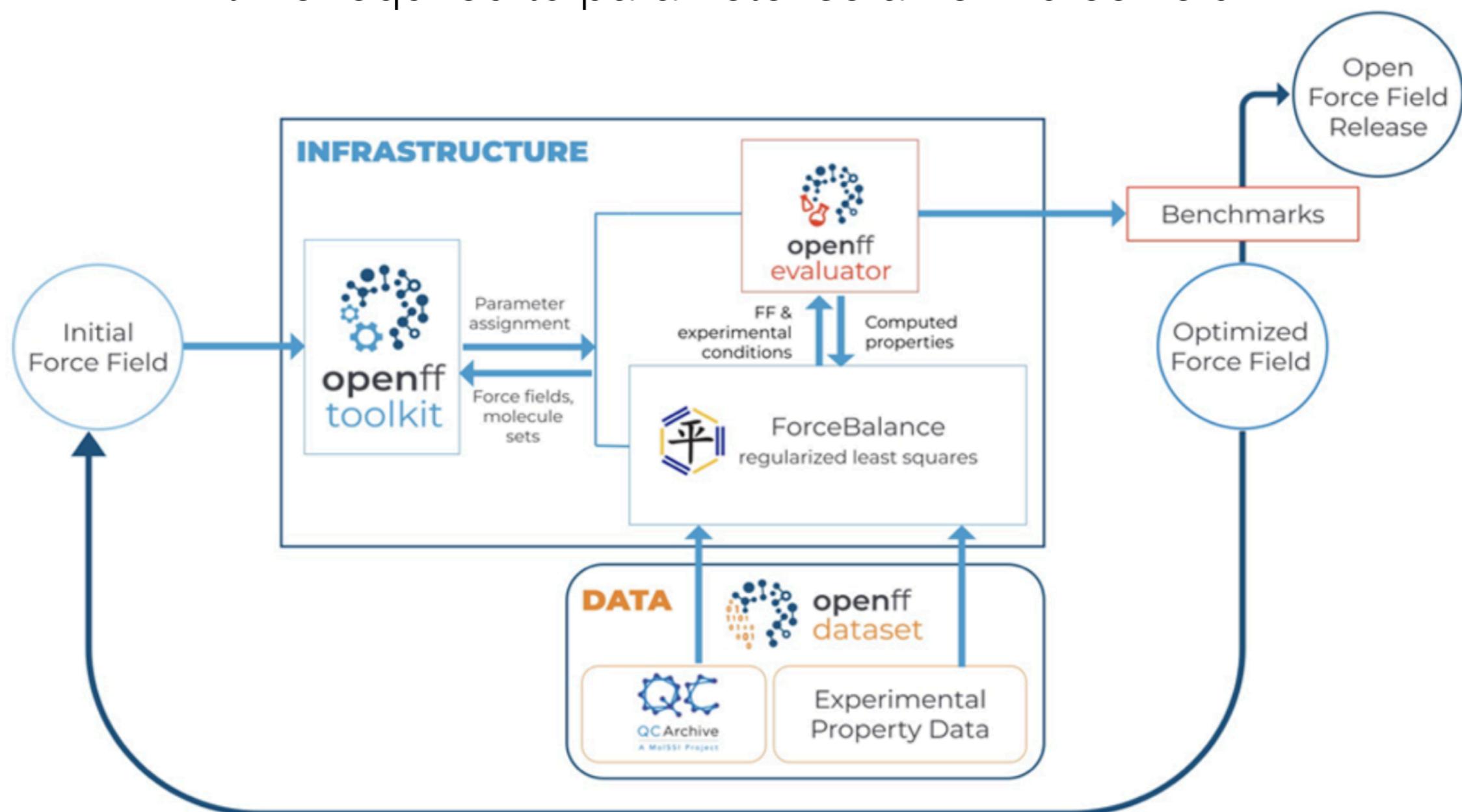
<https://openforcefield.org/>

# Open Force Field Timeline



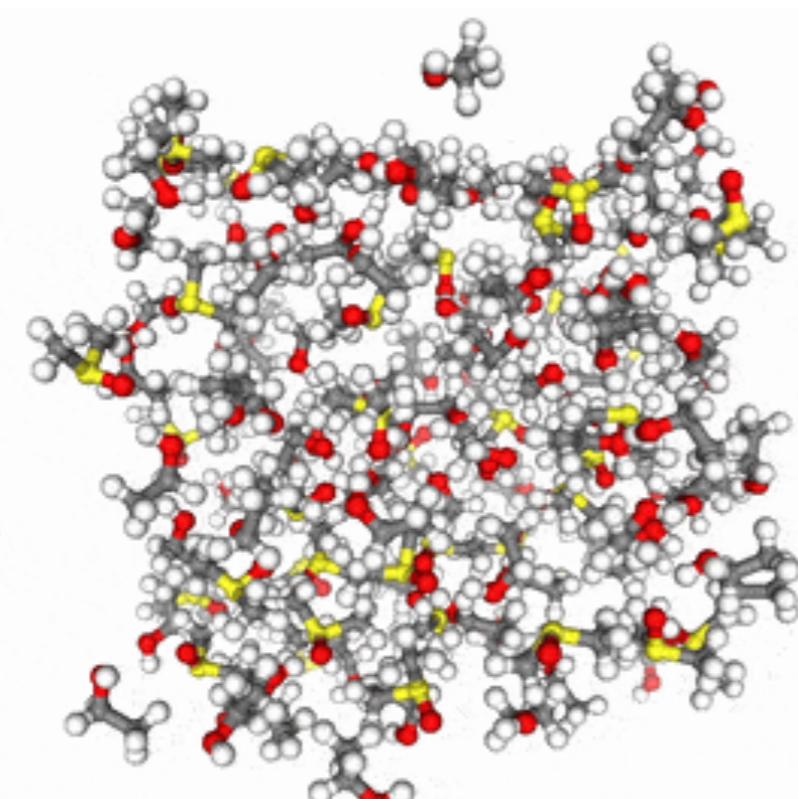
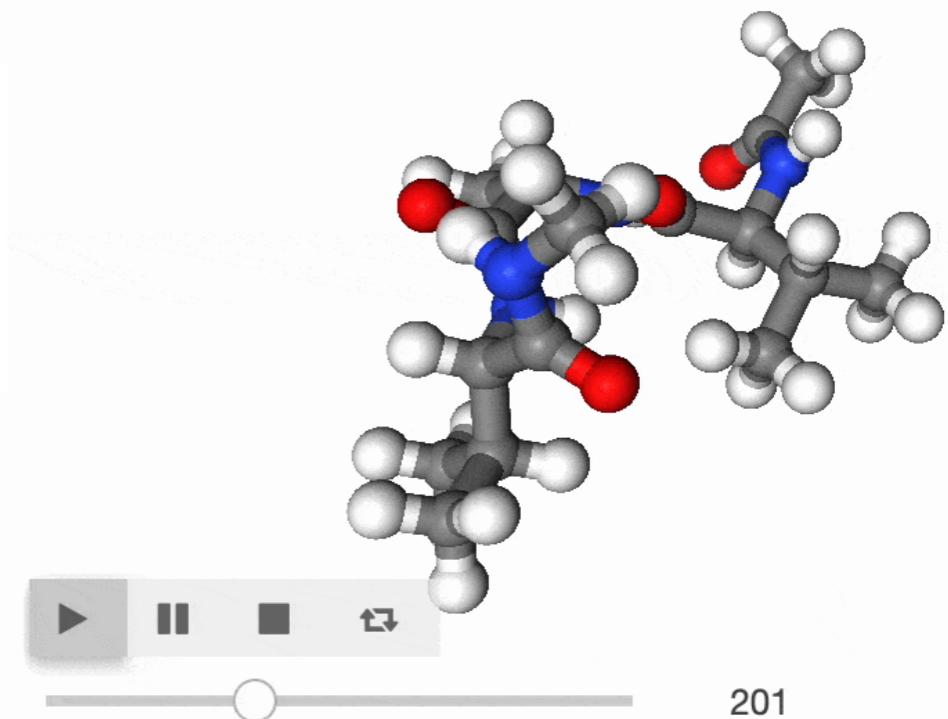
# Open Force Field

Open Force Field infrastructure significantly reduces human time required to parameterise a new force field.



# Training Data

Bonded parameters trained  
on curated QM energies and  
forces of 1000's of drug-like  
molecules and fragments

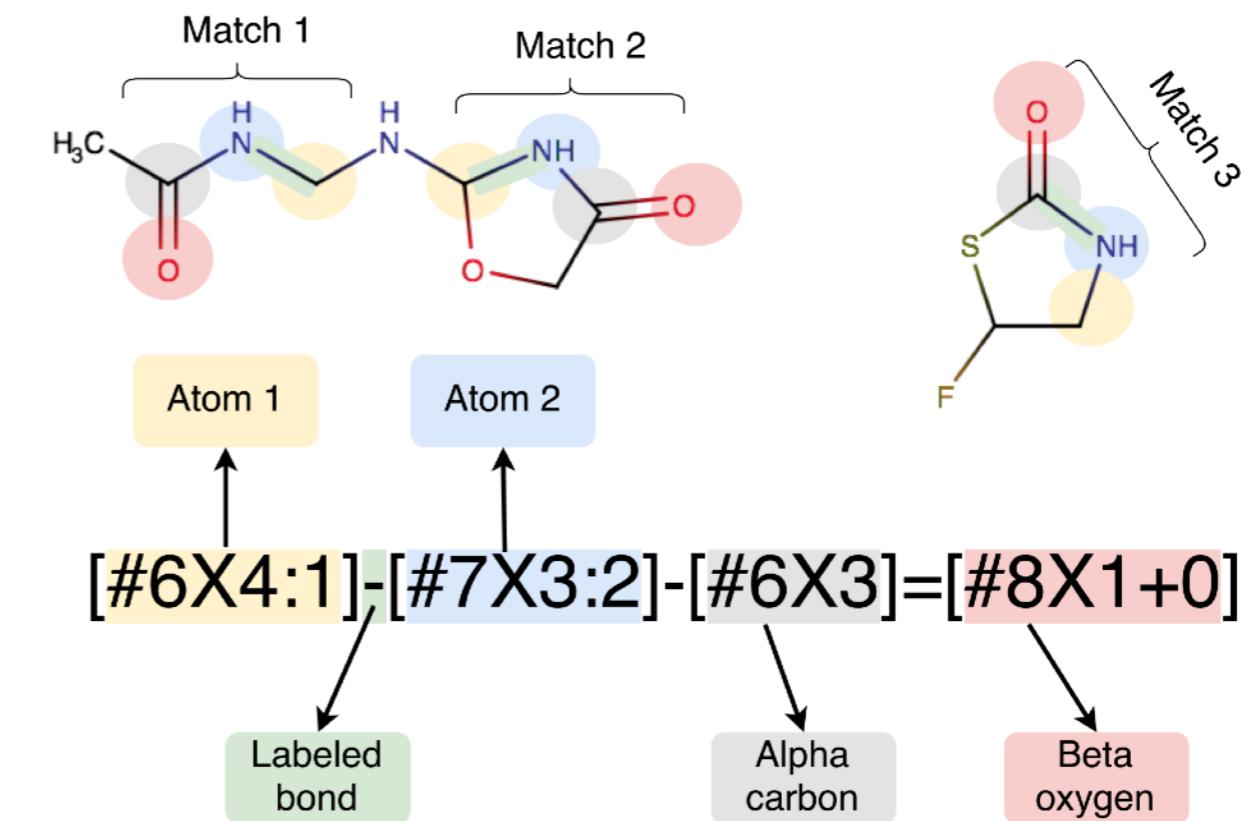


Non-bonded training set  
includes 1000's of mixture  
enthalpy and density  
measurements from NIST  
ThermoML

# SMIRKS Native Open Force Field Specification

Direct chemical perception bypasses atom typing and assigns parameters to atoms/bonds/angles/torsions by processing the full molecular graph.

Mobley et al., JCTC, 2018, **14**, 6076



Can avoid a proliferation of parameters without loss of accuracy.

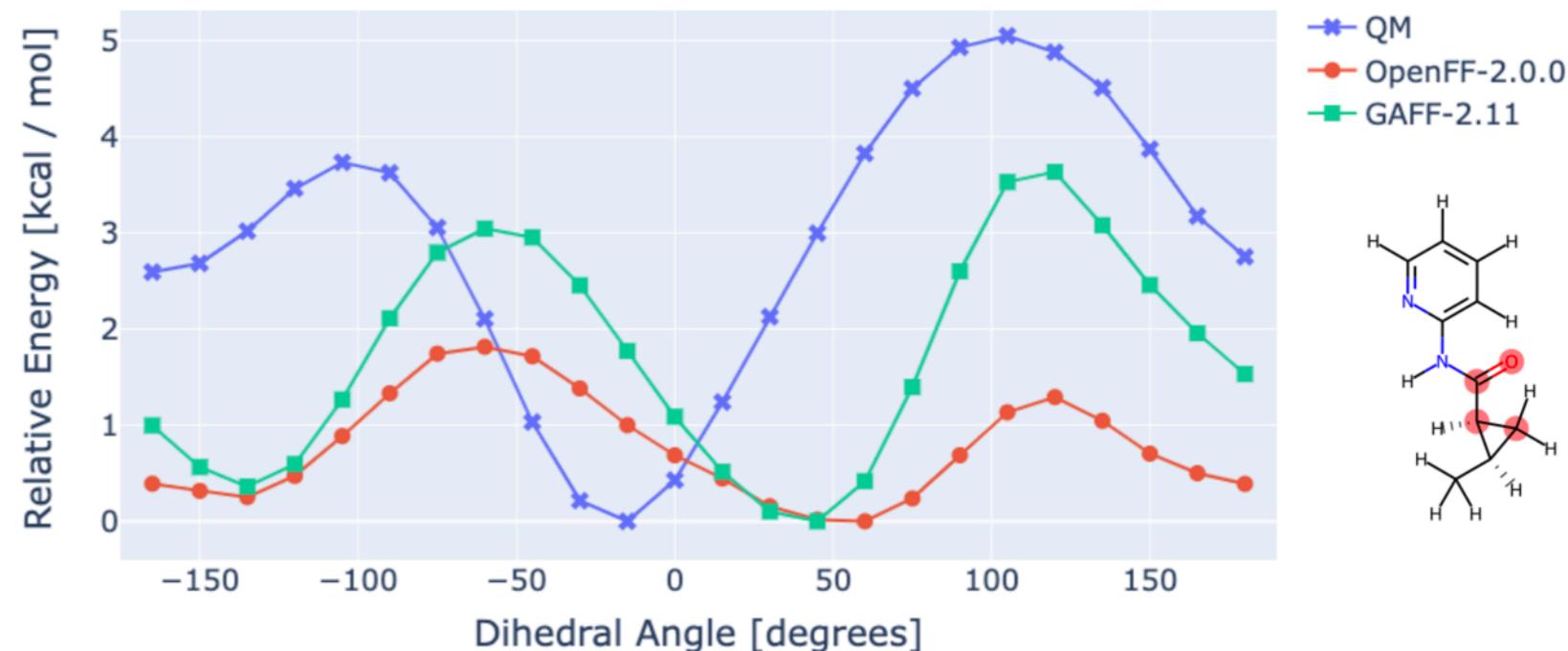
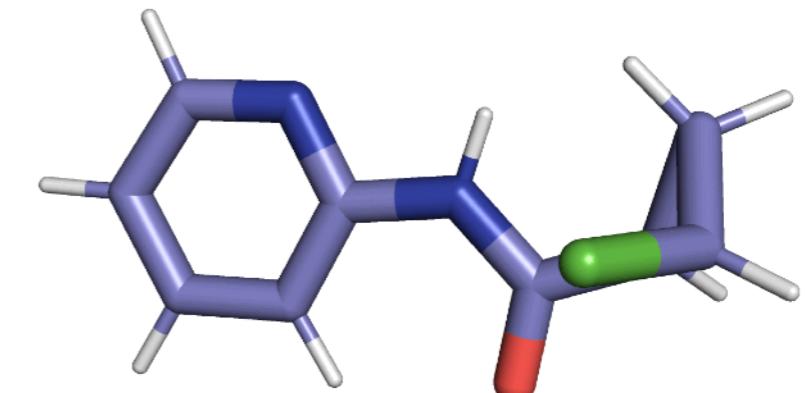
| parameter type     | MMFF <sup>5</sup> | OPLS3 <sup>5</sup> | OPLS3e <sup>4</sup> | Sage (OpenFF 2.0.0) <sup>6</sup> |
|--------------------|-------------------|--------------------|---------------------|----------------------------------|
| bond stretching    | 456               | 1187               | 1187                | 88                               |
| angle bending      | 2283              | 15,236             | 15,235              | 40                               |
| torsional rotation | 520               | 48,142             | 146,669             | 167                              |

# Open Force Field BespokeFit

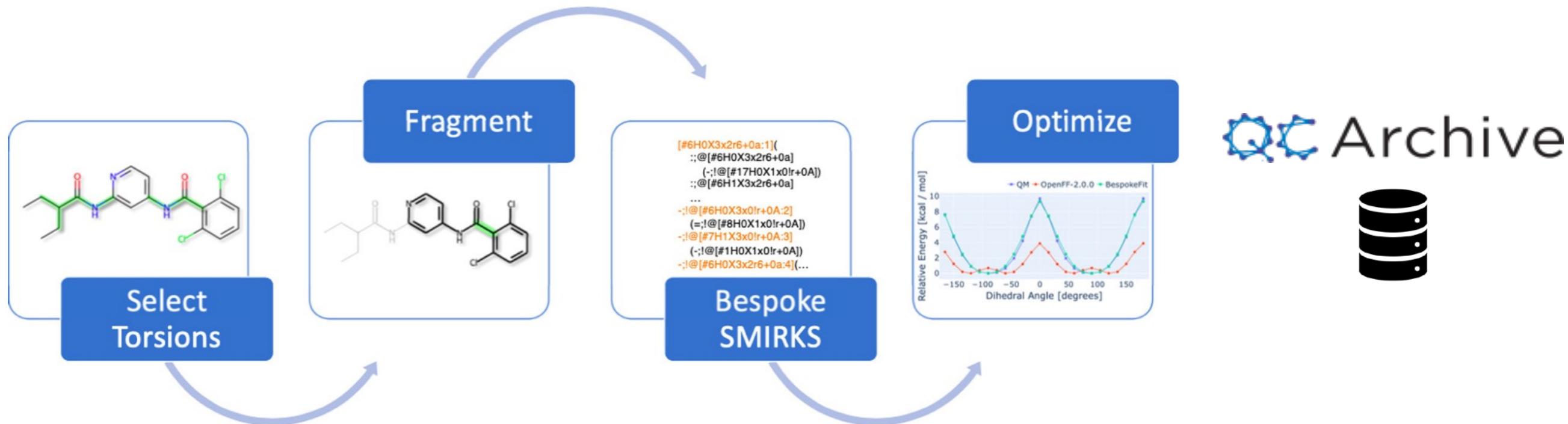
Accurate determination of molecular conformation is crucial in structure-based drug design.

Molecular conformation is largely determined by torsional rotation about flexible bonds.

Transferability is difficult due to sensitivity to surrounding environment.

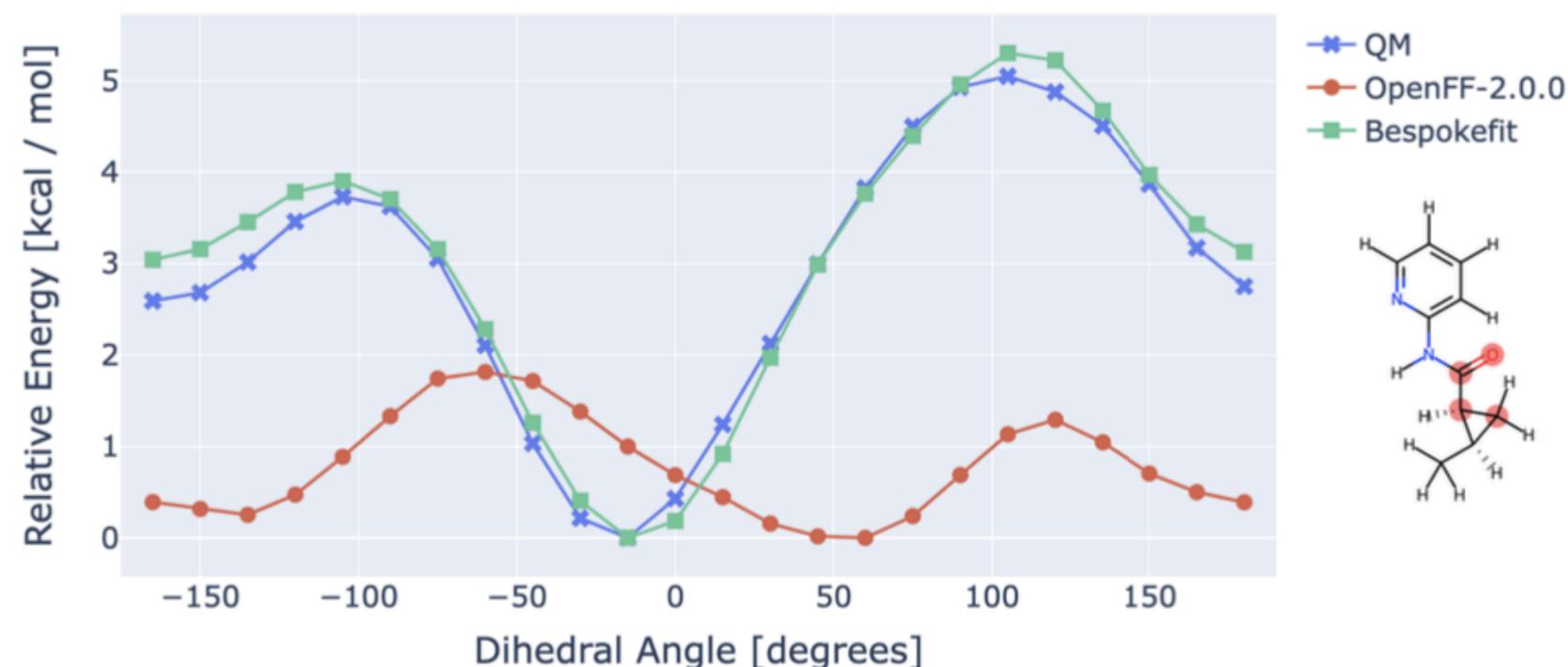


# Open Force Field BespokeFit



**OpenFF-BespokeFit** provides robust molecule-specific parameterization workflow.

Can generate parameters at-scale from community generated data stored on QCArchive.



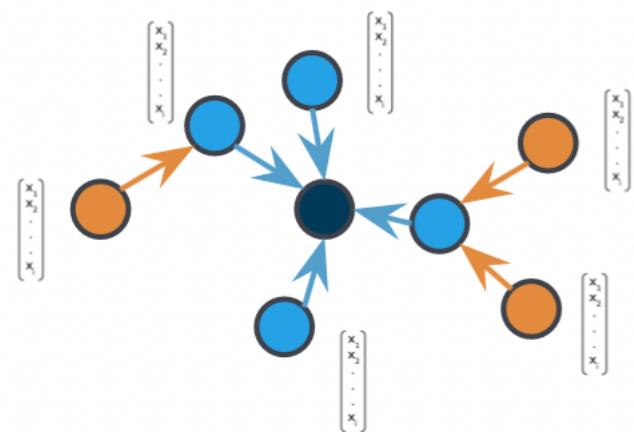
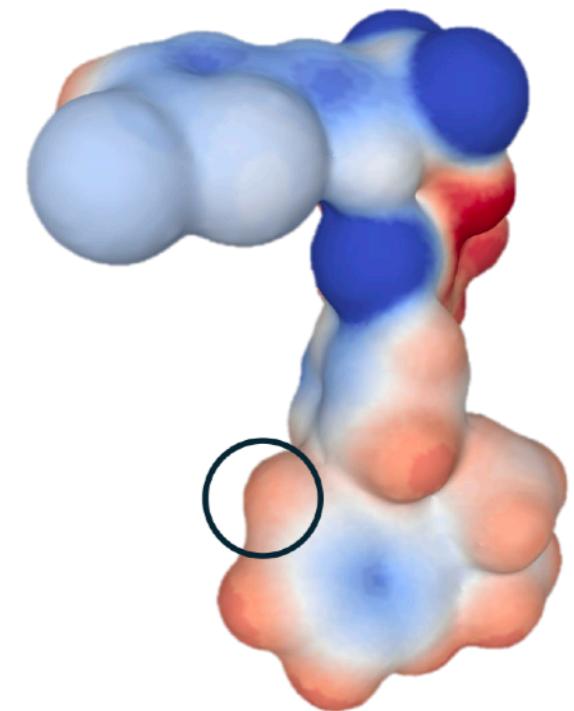
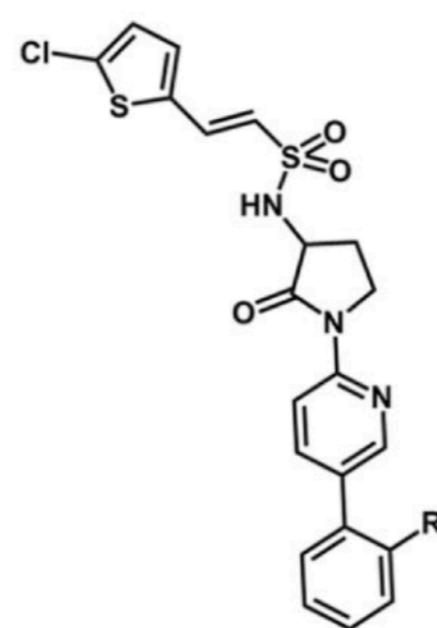
<https://github.com/openforcefield/openff-bespokefit>

Horton et al. JCIM 2022

# Atomic Charges

Most parameters are read in from a library,  
but exception are the atomic charges.

We typically want the charges to reproduce  
the electrostatic potential (ESP) around the  
molecule. (While accounting for condensed  
phase polarisation, ask me if interested!)



Example charge model is RESP, fit to HF/6-31G(d) QM.

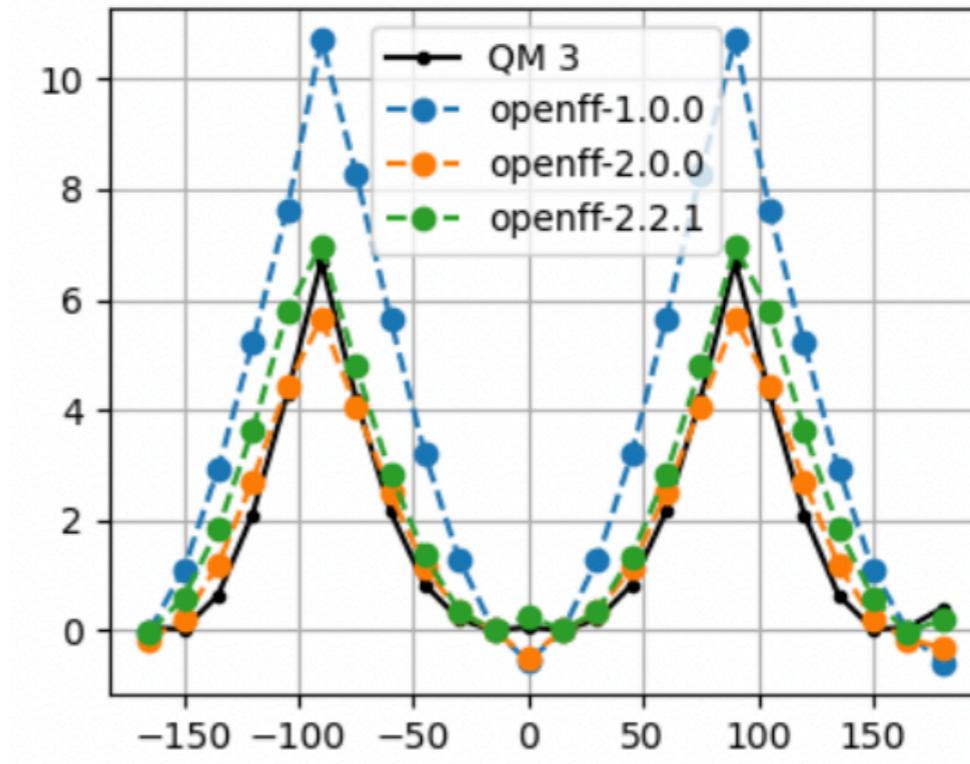
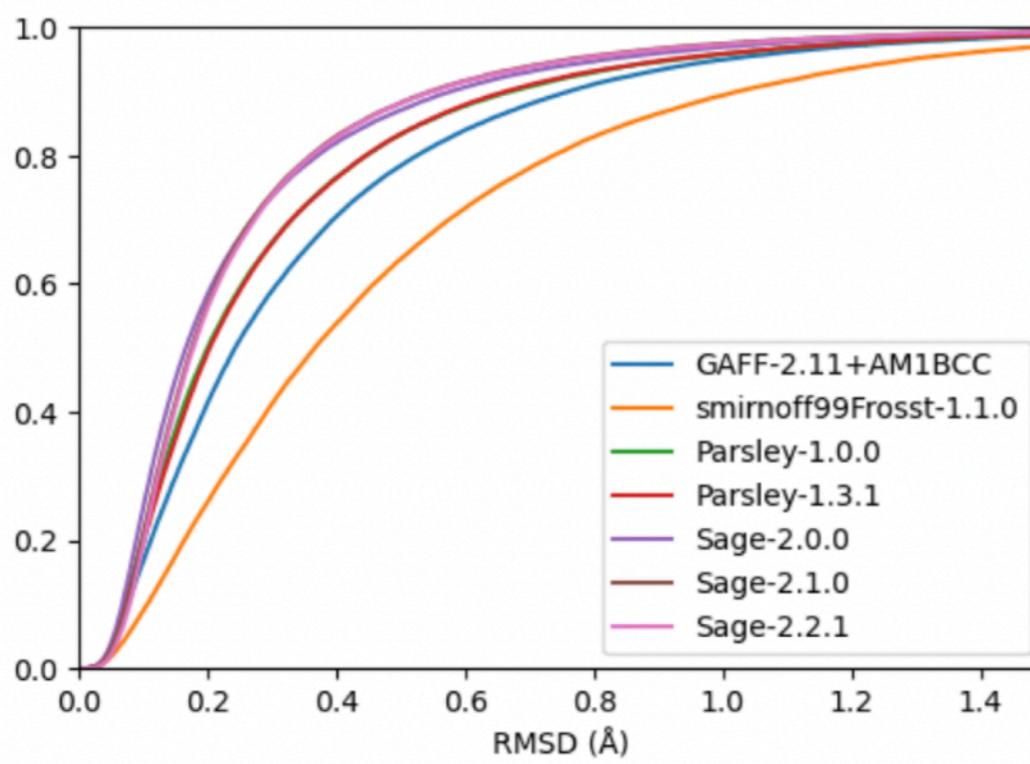
This can be expensive for high-throughput calculations, so alternative is AM1-BCC based on semi-empirical calculations (method used in OpenFF).

This can still be expensive! So a new alternative is graph-based charges (AshGC — see Workshop)

# Infrastructure

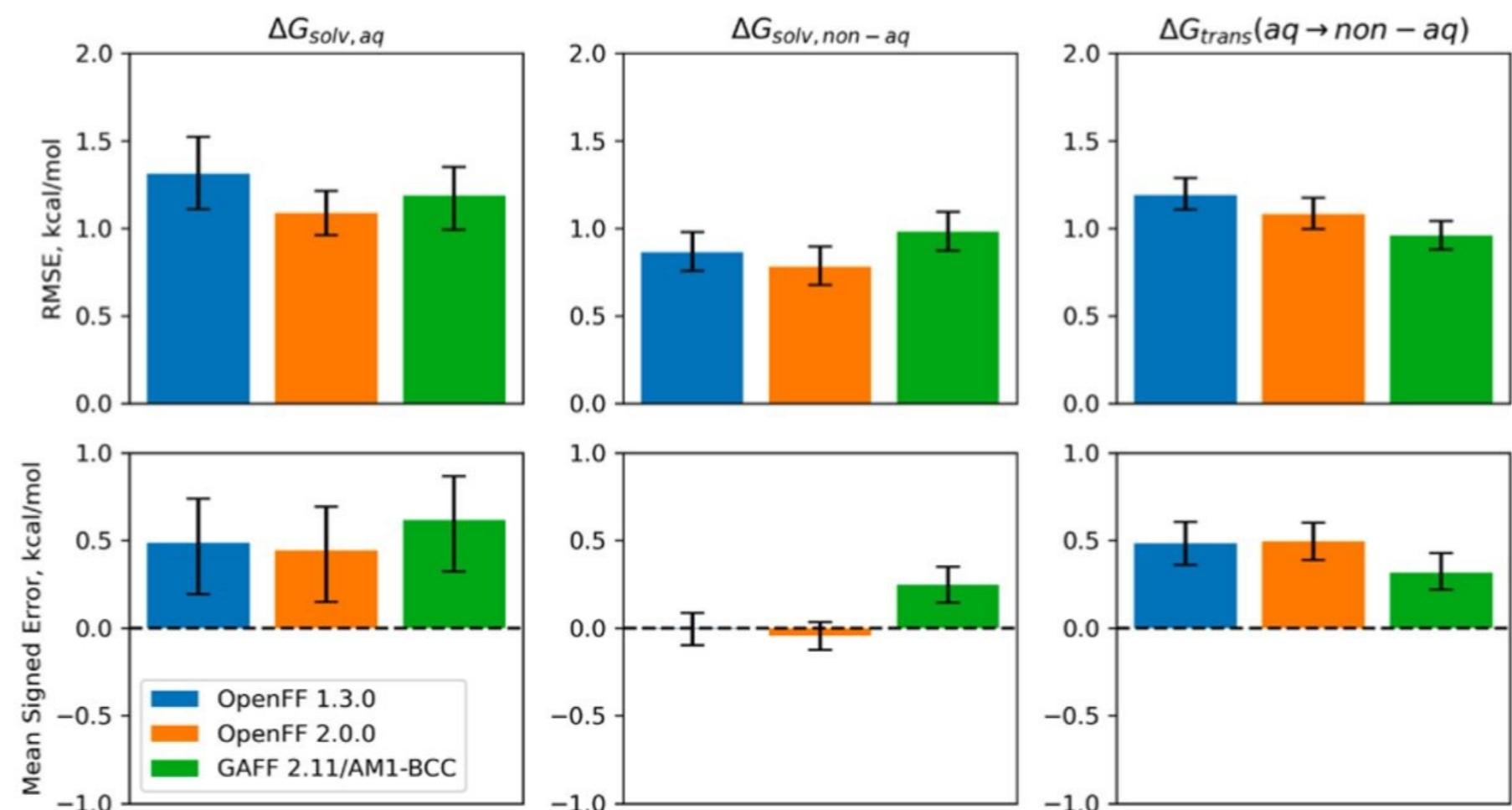
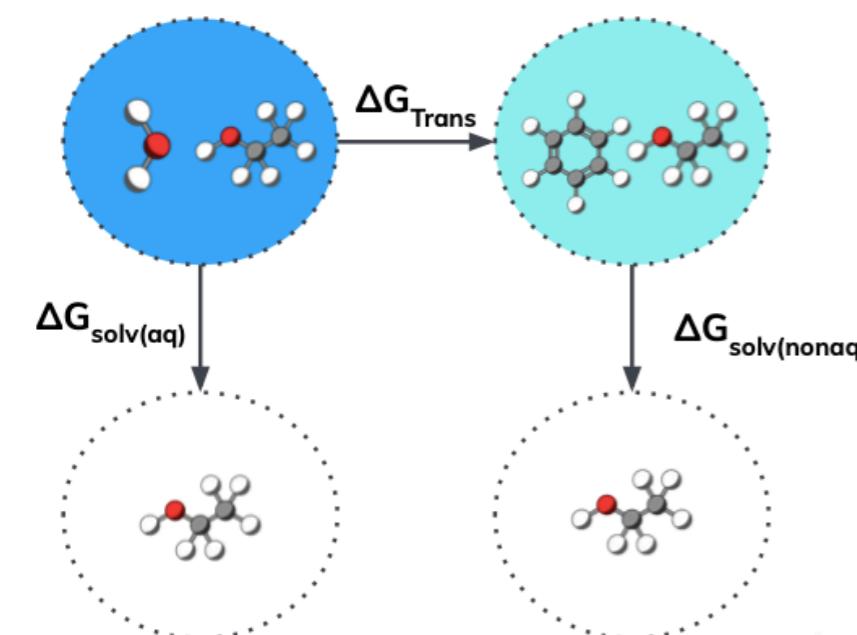
Makes it easy to compare force field versions.

E.g. RMSD between QM and MM minima for QM datasets.



# Infrastructure

Improved performance for solvation / transfer free energies



# Some Current OpenFF Projects

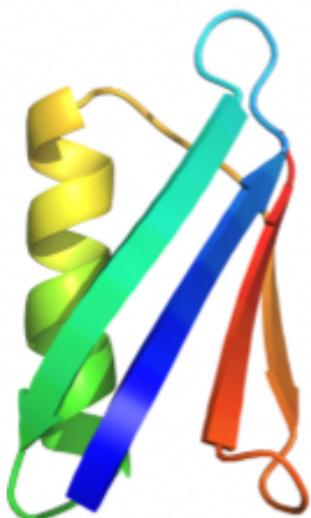
Powered by:



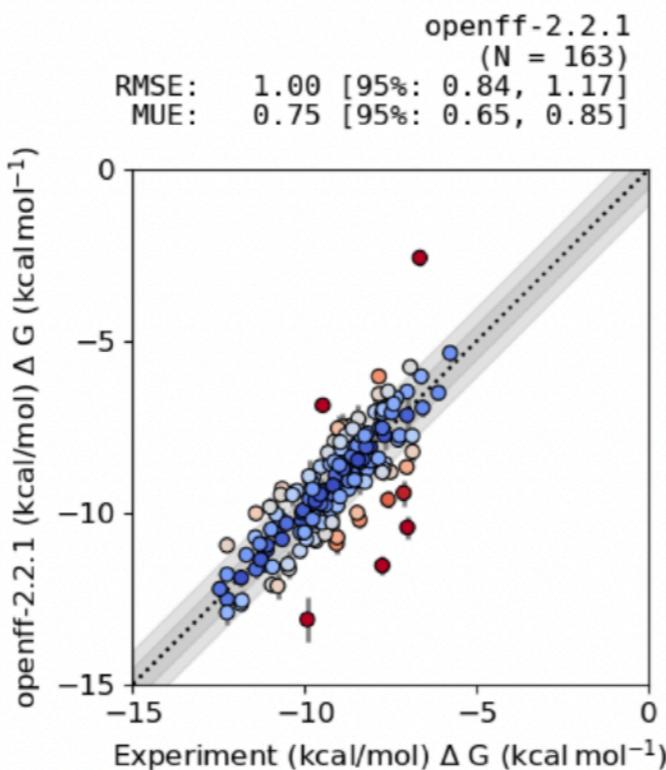
OpenFE Project Staff



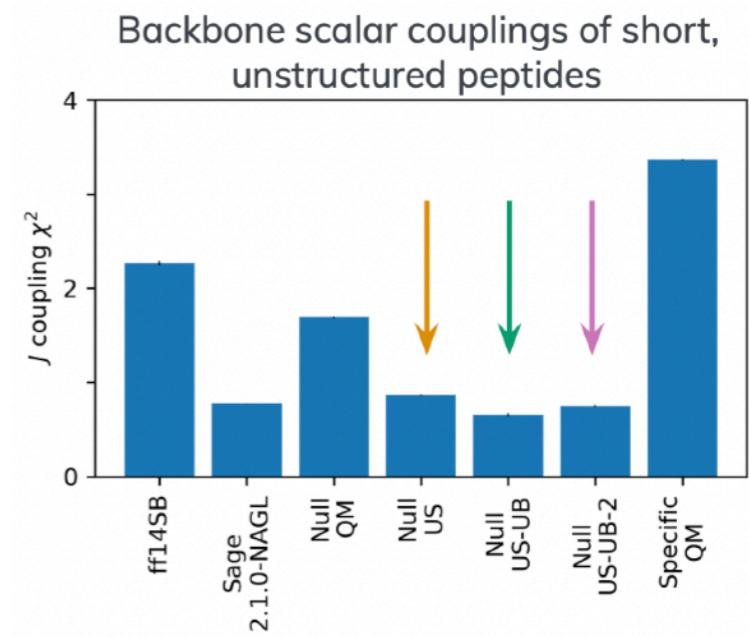
NRP NATIONAL RESEARCH  
PLATFORM



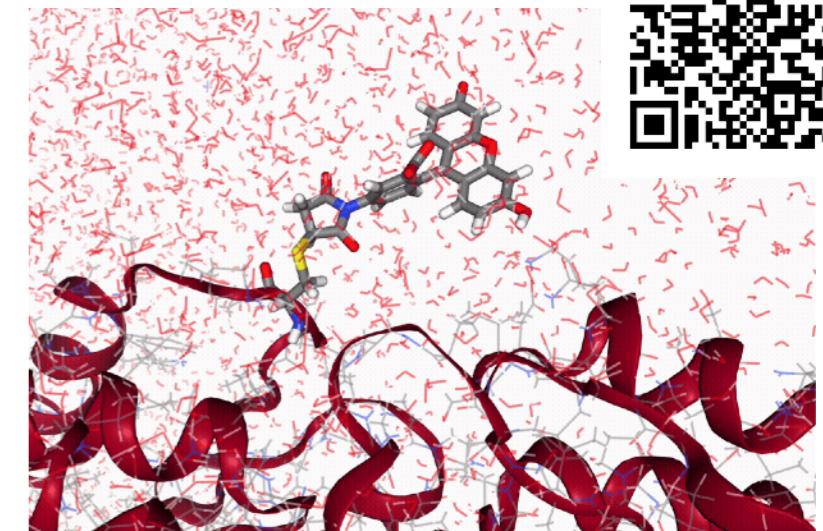
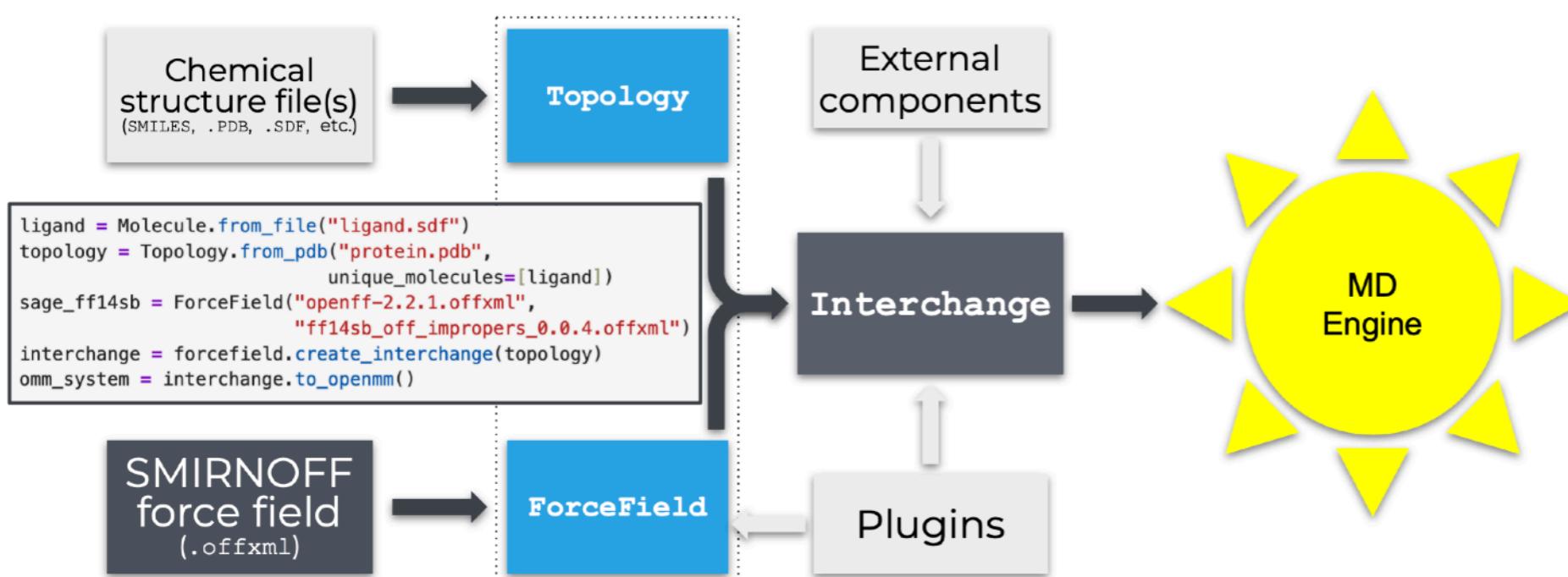
Development of compatible protein force field is underway. Iterative fitting to NMR observables improves stability of MD simulations without folding everything into  $\alpha$ -helix.



Protein-ligand binding free energy benchmarks run at-scale (including support for bespoke parameters)



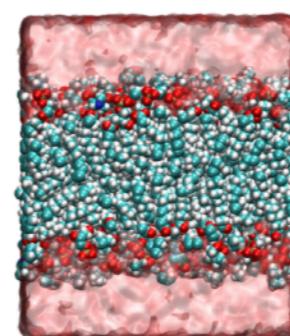
# Some Current OpenFF Projects



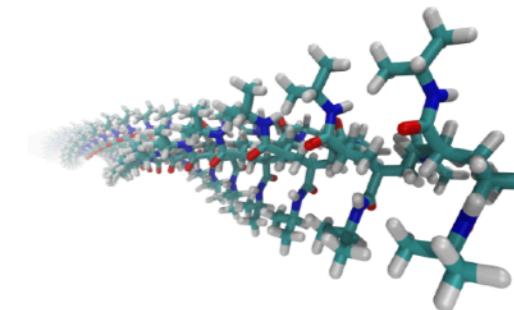
The OpenFF workflow lets you set up MD simulations in a few lines of code

Plus work on lipids, polymers, water, ions,

...



lipids



polymers

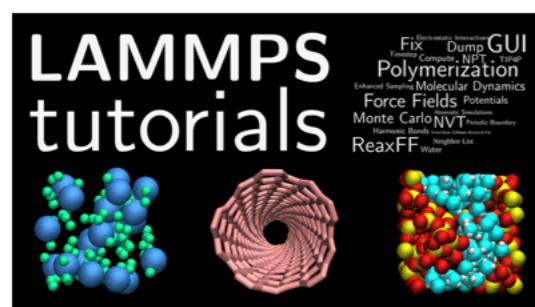
# Where can I find out more?

## Best Practices for Foundations in Molecular Simulations [Article v1.0]

Efrem Braun<sup>1</sup>, Justin Gilmer<sup>2</sup>, Heather B. Mayes<sup>3</sup>, David L. Mobley<sup>4</sup>, Jacob I. Monroe<sup>5</sup>, Samarjeet Prasad<sup>6</sup>, Daniel M. Zuckerman<sup>7</sup>

## A suite of tutorials for the BioSimSpace framework for interoperable biomolecular simulation [Article v1.0]

Lester O. Hedges<sup>1,2\*</sup>, Sofia Bariami<sup>3†</sup>, Matthew Burman<sup>2</sup>, Finlay Clark<sup>3</sup>, Benjamin P. Cossins<sup>4</sup>, Adele Hardie<sup>3</sup>, Anna M. Herz<sup>3</sup>, Dominykas Lukauskis<sup>5</sup>, Antonia S.J.S. Mey<sup>3</sup>, Julien Michel<sup>2,3\*</sup>, Jenke Scheen<sup>3‡</sup>, Miroslav Suruzhon<sup>4</sup>, Christopher J. Woods<sup>1</sup>, Zhiyi Wu<sup>4</sup>



## A Set of Tutorials for the LAMMPS Simulation Package [Article v1.0]

Simon Gravelle, Cecilia M. S. Alvares, Jacob R. Gissinger, Axel Kohlmeyer

[Article Code Repository](#)

[PDF](#)



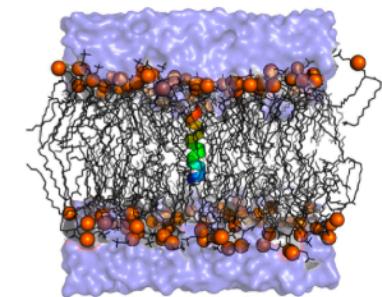
## From Proteins to Perturbed Hamiltonians: A Suite of Tutorials for the GROMACS-2018 Molecular Simulation Package [Article v1.0]

Justin A. Lemkul

Department of Biochemistry, Virginia Polytechnic Institute and State University  
<https://orcid.org/0000-0001-6661-8653>

DOI: <https://doi.org/10.33011/livecoms.1.1.5068>

Keywords: tutorials, gromacs, molecular dynamics simulation, computational chemistry



## Molecular Dynamics: From Basics to Application [Article v1.0]

Luis Vollmers, Shu-Yu Chen, Maria Reif, Tristan Alexander Mauck, Martin Zacharias

[PDF](#)

[Article Code Repository](#)



# Living Journal of Computational Molecular Science

Living Journal of Computational Molecular Science is a diamond open access journal sharing best practices in molecular simulation.

**Living documents**, article types include:

- Best practices
- Tutorials and training
- Software analyses
- Perpetual reviews
- Lessons learned

Send in pre-submission letter to discuss your plans

<https://livecomsjournal.org/>

## Plan for the rest of the day

**12-1pm:** Ligand parameterisation and charge assignment (including AshGC)

**2-3pm:** Running MD and analysis of a protein-ligand system

**3.30-5pm:** Introduction to more simulation analysis methods

