The Open Force Field Consortium 2020: Current overview and future outlook

David L. Mobley[1,2], John D. Chodera[4], Michael R. Shirts[5], Michael K. Gilson[6], Lee-Ping Wang[7], Jeffrey R. Wagner[8], Karmen Condić-Jurkić[4]

[1]Department of Pharmaceutical Sciences, University of California, Irvine CA 92697; [2]Department of Chemistry, University of California, Irvine CA 92697; [4]Computational and Systems Biology Program, Memorial Sloan Kettering Cancer Center, New York NY 10065; [5]Department of Chemical and Biological Engineering, University of Colorado Boulder, Boulder, CO 80309; [6]Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California, San Diego; [7]Department of Chemistry, University of California, Davis; [8]Department of Pharmaceutical Sciences, University of California, Irvine CA 92697

***For correspondence:** info@openforcefield.org (KCJ)

Why new force fields? Molecular modeling is widely used in diverse pharmaceutical discovery applications, but its utility and predictive power is limited by the accuracy and generality of the underlying *molecular mechanics force field* used to compute the energetics of biomolecular systems. Current molecular mechanics force fields are built on early modeling work largely performed in the 1980s and 1990s, and the software infrastructure for developing, refining, and applying them has remained largely unchanged since. While the current models and tools play a key role in the pharmaceutical discovery pipeline, improved accuracy is needed to more effectively guide pharmaceutical discovery and design and to reduce costs and time to market. In summary, a new approach and a modern software infrastructure are needed to allow better force fields to be systematically and rapidly built, updated, and applied.

How we build new force fields The *Open Force Field Consortium* (OpenFF) is an open, industry-funded, pre-competitive effort to build iteratively more accurate force fields to improve predictive design, along with the necessary infrastructure to drive further force field improvements and science. We are taking a three-pronged approach:

- **Open science:** Performing new research to inform design and support development of modern force fields, which is rapidly disseminated through the Initiative's website, events, scientific publications, and social media.
- **Open source:** Building a maintainable, extensible automated Python infrastructure to enable rapid cycles of optimization and facile integration into existing workflows, following the best practices of software development. The core infrastructure components used in force field parameterization are OpenFF Toolkit, ForceBalance, OpenFF Evaluator and QCArchive (see below). These are made publicly available under open source licenses to facilitate replication, reuse, and extension by other research groups.
- **Open data:** Creating and curating datasets used in force field parameterization and benchmarking, released under Creative Commons or similar licenses for easy reuse.

The OpenFF force fields are iteratively improved through automated optimization cycles, and each release is evaluated with an ever-growing benchmark suite. Early feasibility studies identify whether new research innovations (such as off-center charges) are ready to incorporate into automated fitting cycles. Improvements may derive, for example, from better data selection protocols, optimization procedures, type definitions, or functional forms. All software, documentation, data, and force fields are made openly available, as detailed at https://openforcefield.org/, allowing industry partners to make use of our force fields and infrastructure in their own discovery programs and workflows, enabling other researchers to further test and build on our research products, and ensuring long-term availability.

Progress in the first two years.

Since its launch in Oct 2018, OpenFF has met all milestones to date, which include infrastructure deliverables and force field releases (Figure 1). The first year focused on open toolkit development and automated parameterization and benchmarking workflows. In the second year, we have continued to develop OpenFF infrastructure by adding new features to the existing packages, starting some new projects (bespoke torsion parameterization, system object development, and others) and improving documentation. In addition, we focused on improving our force fields through a series of targeted scientific studies aimed at devising better performing chemical perception and data selection algorithms, and benchmarking workflows.

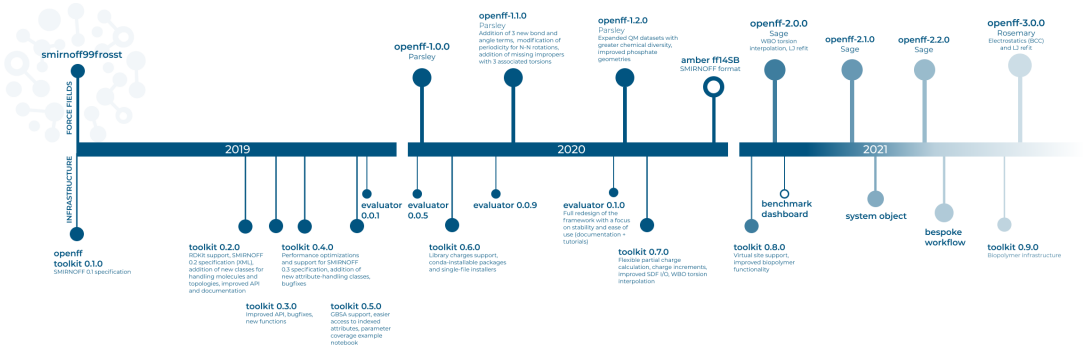

Figure 1. Open Force Field milestones. This graphic highlights the most important achievements of the Open Force Field Consortium since its launch in Oct 2018, and the anticipated progress in Year 3 (2021). All currently available and upcoming force field releases are depicted above the timeline, while the infrastructure progress is shown below.

Some of the most important milestones achieved so far are listed below:

- **Force fields.** We have released three optimized versions of the first generation of Open Force Fields (codename "Parsley") – OpenFF-1.0.0 (10/2019) [1], OpenFF-1.1.0 (03/2020), and OpenFF-1.2.0 (06/2020). After releasing an initial general-purpose force field with unprecedented coverage per number of parameters (SMIRNOFF99Frosst), OpenFF iterated through three refinements of chemical typing, each refit improving accuracy for targeted chemistries (Figure 2 A, B) and addressing other limitations to achieve comparable performance to other publicly available or commercial force fields (Figure 2 C, D).

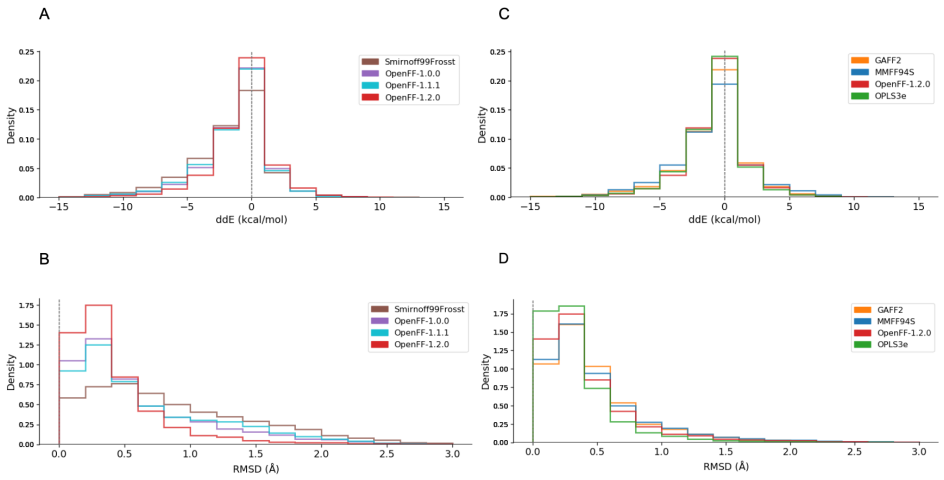

Figure 2. Force field benchmarking demonstrates significant improvements with each generation. Preliminary assessments of OpenFF and other small molecule force fields (FF) compared to quantum mechanical (B3LYP-D3BJ/DZVP) gas phase energies (A, C) and geometries (B,D) of a selected set of molecular conformers.

Our most recent release, OpenFF-1.2.0, seems to perform particularly well for test datasets that include industry-provided molecules, such as a set of particularly challenging fragments from Pfizer. Furthermore, as recently reported[2], we obtained promising results in relative binding free energy calculations for a congeneric series of ligands of the non-receptor tyrosine kinase Tyk2, where OpenFF-1.0.0 was used for ligands and AMBER14SB and TIP3P for protein and water, respectively (Figure 3). The calculated RMSE of 0.97 [95% CI: 0.68, 1.22] kcal/mol is statistically indistinguishable from the Schrödinger JACS benchmarking result [3].

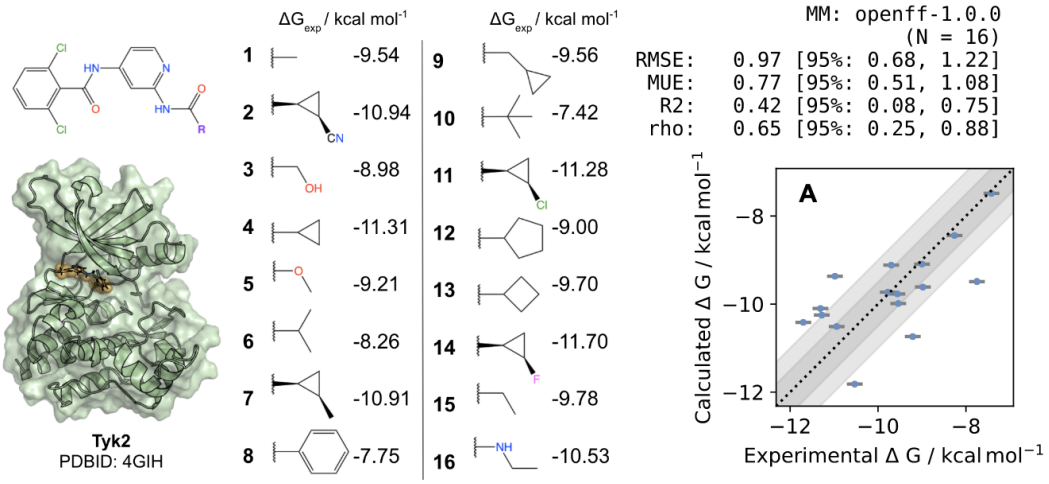

Figure 3. Force field accuracy approaching OPLS3e. Relative alchemical free energy calculations on the Tyk2 kinase:inhibitor system from the Schrödinger JACS benchmark set show OpenFF 1.0.0 achieves an RMSE of 0.97 [95% CI: 0.68, 1.22] kcal/mol, statistically indistinguishable from the results reported by Schrödinger [2].

- **Software.** Most of the early efforts of the OpenFF Initiative were directed toward building an automated infrastructure for force field optimization and benchmarking (Figure 4). All our software can be found on GitHub. The core software development projects include:
  - **OpenFF Toolkit**: a Python toolkit providing implementations of OpenFF's innovative force field specification (SMIRNOFF), parameterization engine, and many other tools. Some of the capabilities include charge assignment (via library charge or flexible partial charge calculations), torsion interpolation based on Wiberg Bond Order, GBSA support, parameter coverage checks, etc. Seven increasingly powerful versions of the toolkit have been released up to date. Full documentation can be found here.
  - **OpenFF Evaluator**: an automated and scalable framework for curating, manipulating, and computing datasets of physical properties from molecular simulation and experimental data. This flexible framework has added new features with each release, and one of the latest versions (v0.1.0) constitutes almost a full redesign of the framework with a focus on stability and ease of use. Full documentation can be found here.
  - **ForceBalance**: a powerful and highly versatile software package for force field optimization. Integration of ForceBalance with the OpenFF Toolkit and OpenFF Evaluator was an important step in development of our fully automated force field optimization infrastructure.
  - **QCArchive**: a free, community-driven, multi-user quantum chemistry database. Adaptation and use of QCArchive components to compute and store quantum chemistry data used in force field parameterization and benchmarking has been a key part of OpenFF infrastructure development.

  More information about the anticipated infrastructure projects and activities can be found on the current OpenFF infrastrucure roadmap.

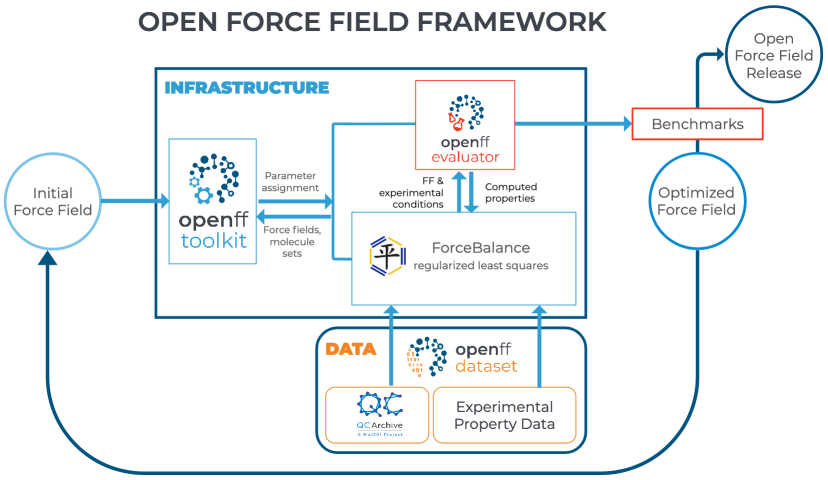

Figure 4. Open Force Field framework. The OpenFF software framework used in force field optimization and benchmarking cycles.

- **Research.** To this date, OpenFF research has led to publication of six peer-reviewed articles [4–9] and two preprints [10, 11] that are headed toward peer-reviewed publication. The full list of publications and preprints that resulted from the Open Force Field Consortium/Initiative's work to date is available on our website. Our scientific roadmap provides a full list of planned research projects for this year and beyond. These projects will be integrated with the broader infrastructure as they mature and test out, and the results of the feasibility studies will be reported early on our website and in preprints, and ultimately in peer-reviewed journals.

Year 3 and beyond. OpenFF research activities are tightly coupled with infrastructure development, and aimed at continuous improvement of the produced force fields. In the coming year, OpenFF will reap major accuracy benefits resulting from infrastructure and science investments made in the first two years. Particularly strong progress is expected in the following areas:

- **Bespoke parameterization workflow** to refit torsions to QM or quantum machine learning (including ANI-2x [12]).
- **Improved torsion accuracy** for the general force field based on Wiberg bond order (WBO)-based parameter interpolation, planned to be included in the OpenFF-2.0.0 force field release (codename "Sage").
- **Improved charge models** via support for AM1-BCC and off-center charges, and machine learning-based charge models. A feasibility study will test which model provides the best balance between speed and accuracy, and that model will be subsequently used for a charge refit in the OpenFF-3.0.0 force field generation (codename "Rosemary").
- **Automated typing inference** (for selected cases) and continued improvements to chemical perception to enhance force field performance, potentially allowing fit of full force field from scratch and removing any legacy problems in the force fields.
- **Extensive, ongoing assessment of force field performance** with respect to QM conformational energetics and to experimental protein-ligand binding free energies, with comparisons against other publicly available force fields, using automated benchmarking protocols and infrastructure.
- **Consistent small molecule, biopolymer, and solvent force fields** leading to accuracy improvements. This component benefits from synergy with NIH support of the OpenFF Initiative's effort to produce a consistent biopolymer force field.

Additional scientific innovations will be integrated to improve accuracy as parallel science efforts demonstrate their utility in protein-ligand modeling. An additional long-term research and infrastructure aim in-

cludes development and implementation of Bayesian inference and surrogate modeling as a central part of our force field optimization infrastructure, also supported by the NIH grant.

References

[1] Yudong Qiu, Daniel G. A. Smith, Simon Boothroyd, Jeffrey Wagner, Caitlin C. Bannan, Trevor Gokey, Hyesu Jang, Victoria T. Lim, Chaya D. Stern, Andrea Rizzi, Xavier Lucas, Bryon Tjanaka, Michael R. Shirts, Michael K. Gilson, John D. Chodera, Christopher I. Bayly, David L. Mobley, and Lee-Ping Wang. Introducing the first optimized Open Force Field 1.0.0 (codename "Parsley"), October 2019. URL https://doi.org/10.5281/zenodo.3963295.

[2] Dominic A. Rufa, Hannah E. Bruce Macdonald, Josh Fass, Marcus Wieder, Patrick B. Grinaway, Adrian E. Roitberg, Olexandr Isayev, and John D. Chodera. Towards chemical accuracy for alchemical free energy calculations with hybrid physics-based machine learning / molecular mechanics potentials. *bioRxiv*, 2020. doi: 10.1101/2020.07.29.227959. URL https://www.biorxiv.org/content/early/2020/07/30/2020.07.29.227959.

[3] Lingle Wang, Yujie Wu, Yuqing Deng, Byungchan Kim, Levi Pierce, Goran Krilov, Dmitry Lupyan, Shaughnessy Robinson, Markus K. Dahlgren, Jeremy Greenwood, Donna L. Romero, Craig Masse, Jennifer L. Knight, Thomas Steinbrecher, Thijs Beuming, Wolfgang Damm, Ed Harder, Woody Sherman, Mark Brewer, Ron Wester, Mark Murcko, Leah Frye, Ramy Farid, Teng Lin, David L. Mobley, William L. Jorgensen, Bruce J. Berne, Richard A. Friesner, and Robert Abel. Accurate and reliable prediction of relative ligand binding potency in prospective drug discovery by way of a modern free-energy calculation protocol and force field. *Journal of the American Chemical Society*, 137(7): 2695–2703, 2015. doi: 10.1021/ja512751q. URL https://doi.org/10.1021/ja512751q. PMID: 25625324.

[4] David L. Mobley, Caitlin C. Bannan, Andrea Rizzi, Christopher I. Bayly, John D. Chodera, Victoria T. Lim, Nathan M. Lim, Kyle A. Beauchamp, David R. Slochower, Michael R. Shirts, Michael K. Gilson, and Peter K. Eastman. Escaping atom types in force fields using direct chemical perception. *Journal of Chemical Theory and Computation*, 14(11):6076–6092, 2018. doi: 10.1021/acs.jctc.8b00640. URL https://doi.org/10.1021/acs.jctc.8b00640. PMID: 30351006.

[5] Camila Zanette, Caitlin C. Bannan, Christopher I. Bayly, Josh Fass, Michael K. Gilson, Michael R. Shirts, John D. Chodera, and David L. Mobley. Toward learned chemical perception of force field typing rules. *Journal of Chemical Theory and Computation*, 15(1):402–423, 2019. doi: 10.1021/acs.jctc.8b00821. URL https://doi.org/10.1021/acs.jctc.8b00821.

[6] K. Kellett, B. M. Duggan, and M. K. Gilson. Facile synthesis of a diverse library of mono-3-substituted beta-cyclodextrin analogues. *Supramolecular Chemistry*, 31(4):251–259, 2019. doi: 10.1080/10610278.2018.1562191. URL https://doi.org/10.1080/10610278.2018.1562191.

[7] Yudong Qiu, Paul S. Nerenberg, Teresa Head-Gordon, and Lee-Ping Wang. Systematic optimization of water models using liquid/vapor surface tension data. *The Journal of Physical Chemistry B*, 123(32):7061–7073, 2019. doi: 10.1021/acs.jpcb.9b05455. URL https://doi.org/10.1021/acs.jpcb.9b05455. PMID: 31314516.

[8] David R. Slochower, Niel M. Henriksen, Lee-Ping Wang, John D. Chodera, David L. Mobley, and Michael K. Gilson. Binding thermodynamics of host–guest systems with smirnoff99frosst 1.0.5 from the open force field initiative. *Journal of Chemical Theory and Computation*, 15(11):6225–6242, 2019. doi: 10.1021/acs.jctc.9b00748. URL https://doi.org/10.1021/acs.jctc.9b00748. PMID: 31603667.

[9] Yudong Qiu, Daniel G. A. Smith, Chaya D. Stern, Mudong Feng, Hyesu Jang, and Lee-Ping Wang. Driving torsion scans with wavefront propagation. *The Journal of Chemical Physics*, 152(24):244116, 2020. doi: 10.1063/5.0009232. URL https://doi.org/10.1063/5.0009232.

[10] Caitlin C. Bannan and David Mobley. Chemper: An open source tool for automatically generating smirks patterns. *ChemRxiv*, 2019. doi: 10.26434/chemrxiv.8304578.v1. URL https://chemrxiv.org/articles/preprint/ChemPer_An_Open_Source_Tool_for_Automatically_Generating_SMIRKS_Patterns/8304578.

[11] Victoria T. Lim and David Mobley. Benchmark assessment of molecular geometries and energies from small molecule force fields. *ChemRxiv*, 2020. doi: 10.26434/chemrxiv.12551867.v1. URL https://chemrxiv.org/articles/preprint/Benchmark_Assessment_of_Molecular_Geometries_and_Energies_from_Small_Molecule_Force_Fields/12551867.

[12] Christian Devereux, Justin Smith, Kate Davis, Kipton Barros, Roman Zubatyuk, Olexandr Isayev, and Adrian Roitberg. Extending the applicability of the ANI deep learning molecular potential to sulfur and halogens. *ChemRxiv*, 2020. doi: 10.26434/chemrxiv.11819268.v1. URL https://chemrxiv.org/articles/preprint/Extending_the_Applicability_of_the_ANI_Deep_Learning_Molecular_Potential_to_Sulfur_and_Halogens/11819268.