The Journal of Open Source Software

MoSDeF-dihedral-fit: A lightweight software for fitting dihedrals within MoSDeF

Brad Crawford [1,2], **Co D. Quach** [3,4], **Nicholas C. Craven** [4,5], **Christopher R. Iacovella** [3,4,5], **Clare McCabe** [3,4,6], **Peter T. Cummings** [3,4,6], **and Jeffrey J. Potoff** [2]

1 Atomfold, PA, United States of America **2** Department of Chemical Engineering and Materials Science, Wayne State University, Detroit, MI 48202-4050, United States of America **3** Department of Chemical and Biomolecular Engineering, Vanderbilt University, Nashville, TN 37235-1604, United States of America **4** Multiscale Modeling and Simulation (MuMS) Center, Vanderbilt University, Nashville, TN 37212, United States of America **5** Interdisciplinary Material Science Program, Vanderbilt University, Nashville, TN 37235-0106, United States of America **6** School of Engineering and Physical Sciences, Heriot-Watt University, Edinburgh EH14 4AS, United Kingdom

DOI: 10.21105/joss.06970

Software

- Review
- Repository
- Archive

Editor: Sophie Beck

Reviewers:

- @naik-aakash
- @RiesBen

Submitted: 17 May 2024
Published: 06 December 2024

Summary

Molecular Mechanics (MM) simulations (e.g., molecular dynamics and Monte Carlo) provide a third method of scientific discovery, adding to the traditional theoretical and experimental scientific methods (Mielke et al., 2019; Siegfried, 2014). Experimental methods measure the data under set conditions (e.g., temperature and pressure), whereas the traditional theoretical methods are based on analytical equations, and sometimes their constants are fitted to experimental data. The MM simulations are deterministic and stochastic, and their models, commonly known as "force fields", can be optimized to match experimental data, similar to analytical theory-based methods (Allen & Tildesley, 2017; Errington & Panagiotopoulos, 1999; Frenkel & Smit, 2002; Hemmen & Gross, 2015; Jorgensen et al., 1996; Martin & Siepmann, 1998; Potoff & Bernard-Brunel, 2009; Scott J. Weiner et al., 1984; S. J. Weiner et al., 1986). In larger, more complex systems, the stochastic simulation's molecules can jump large energy barriers that deterministic simulations may not be able to overcome in a reasonable timeframe, even with modern computing capabilities (Allen & Tildesley, 2017; Frenkel & Smit, 2002). However, deterministic and stochastic systems that provide adequate sampling for calculating a given property can provide critical insights into the system's phase space, which are not obtainable via traditional theoretical and experimental methods. Additionally, molecular simulations provide critical insights from visualizations and by obtaining chemical or material properties that do not currently exist, are not easily attainable (e.g., too expensive or dangerous) by traditional theoretical and experimental methods (Hirst et al., 2014; Hollingsworth & Dror, 2018), or require hard-to-achieve conditions, such as very high pressures and temperatures (Koneru et al., 2022; Kumar et al., 2022; Louie et al., 2021; Swai, 2020; Yu & Pahl, 2023). However, the force field parameters are ideally determined from Quantum Mechanics (QM) simulations or other methods, including the vibrational spectrum and machine learning methods (Friederich et al., 2018; Kania et al., 2021; Mayne et al., 2013; Schmid et al., 2011; Kenno Vanommeslaeghe et al., 2014; Vermeyen et al., 2023). The MM proper dihedrals (i.e., dihedrals) are challenging to obtain if they do not currently exist for the chosen force field, inaccurately scale-up in larger molecules, or misbehave with other moiety combinations, provided some were separately derived using small molecules (Kania et al., 2021; Mayne et al., 2013). While the same QM simulations can fit the dihedrals in most force field types, these dihedrals are not easily transferable between force fields due to the different parameters and formulas, including the combining rules and 1-4 scaling factors. (Chen et al., 2015; Huang & Roux, 2013; K. Vanommeslaeghe et al., 2010; Kenno Vanommeslaeghe et al., 2014).

The `MoSDeF-Dihedral-Fit` (Crawford, Quach, et al., 2023) library allows users to quickly calculate the MM dihedrals directly from the QM simulations for several force fields (OPLS, TraPPE, AMBER, Mie, and Exp6) (Errington & Panagiotopoulos, 1999; Hemmen & Gross, 2015; Jorgensen et al., 1996; Martin & Siepmann, 1998; Potoff & Bernard-Brunel, 2009; Scott J. Weiner et al., 1984; S. J. Weiner et al., 1986). The user simply has to generate or use an existing Molecular Simulation Design Framework (MoSDeF) force field `.xml` file (Cummings et al., 2021; *GMSO*, 2019; Summers et al., 2020; Timalsina, 2022), provide Gaussian 16 `.log` or Gaussian-style QM simulation files that cover the dihedral rotation (typically between 0-360 degrees), and provide the molecular structure information in a `.mol2` format (Frisch et al., 2016). The `MoSDeF-Dihedral-Fit` software uses the QM and MM data to produce the dihedral for the specific force field, fitting the constants for the OPLS dihedral form (equation 1)

$$\begin{aligned} U_{\text{OPLS}} &= \frac{k_0}{2} \\ &+\frac{k_1}{2} * (1 + \cos(\theta)) + \frac{k_2}{2} * (1 - \cos(2 * \theta)) \\ &+\frac{k_3}{2} * (1 + \cos(3 * \theta)) + \frac{k_4}{2} * (1 - \cos(4 * \theta)) \end{aligned} \tag{1}$$

and then analytically converting them to the Ryckaert-Bellemans torsion (equation 2)

$$\begin{aligned} U_{\text{Ryckaert-Bellemans}} &= C_0 \\ &+C_1 * \cos(\psi) + C_2 * \cos(\psi)^2 \\ &+C_3 * \cos(\psi)^3 + C_4 * \cos(\psi)^4 \\ &\text{where: } \psi = \theta - 180^o \end{aligned} \tag{2}$$

and the periodic dihedral forms (equation 3).

$$\begin{aligned} U_{\text{Periodic}} &= K_0 * (1 + \cos(n_0 * \theta - d_0)) \\ &+K_1 * (1 + \cos(n_1 * \theta - d_1)) + K_2 * (1 + \cos(n_2 * \theta - d_2)) \\ &+K_3 * (1 + \cos(n_3 * \theta - d_3)) + K_4 * (1 + \cos(n_4 * \theta) - d_4) \\ &\text{where: } n_0 = 0; n_1 = 1; n_2 = 2; n_3 = 3; n_4 = 4 \\ &d_0 = 90^o; d_1 = 180^o; d_2 = 0^o; d_3 = 180^o; d_4 = 0^o \end{aligned} \tag{3}$$

This analytical conversion from the OPLS dihedral form requires setting the specified parameters in the Ryckaert-Bellemans torsion and periodic dihedral forms (see equations 2 and 3). The software outputs the calculated MM dihedral points, enabling users to fit unsupported dihedral forms, provided the force fields are supported by the MoSDeF, GPU Optimized Monte Carlo (GOMC), MoSDeF-GOMC (Crawford et al., 2022; Crawford, Timalsina, et al., 2023; Crawford, Quach, et al., 2023; Nejahi et al., 2019, 2021), and vmd-python (Betz, 2016) software (a derivative of the VMD software (Humphrey et al., 1996; Stone et al., 2001)).

Statement of need

While many of these MM force field parameters can be transferred between force fields, such as bonds, angles, and improper dihedrals (often referred to as "impropers"), the proper dihedrals (dihedrals) can not be easily transferred due to the different combining rules (arithmetic and geometric) and 1-4 scaling factors (i.e., between the 1st and 4th bonded atoms) that were used in the development of the original parameters (Berthelot, 1898; Good & Hope, 1970; Lorentz, 1881). The accuracy of these dihedral parameters is critical in obtaining the correct molecular conformations and configurations, which are required for understanding and analyzing the system's microstructure and physical properties (e.g., free energies, viscosities, adsorption loading, diffusion constants, and many more).

Some integrated dihedral fitting software currently exists for AMBER (Horton et al., 2022) or CHARMM-style force fields (Mayne et al., 2013), and other software will fit the dihedral constants to the final MM and QM energies, which need to be calculated by other means (Guvench & MacKerell, 1998). However, there is a need for a simple, generalized software package that supports multiple potential functions, imports QM and MM files, automatically reads and organizes the QM data, calculates the MM energies, auto-corrects the dihedral fit to account for multiple instances of the dihedral, and automatically removes the unusable cosine power series combinations due to this symmetry. The `MoSDeF-dihedral-fit` software accomplishes all this and automatically accounts for any of the common combining rules and the 1-4 scaling factors specified via the MoSDeF `.xml` (i.e., force field) files (Cummings et al., 2021; *GMSO*, 2019; Summers et al., 2020; Timalsina, 2022). By allowing the user to set any other dihedral in the molecule to zero, this software avoids forcing one dihedral fit to correct the inaccurate forces of another dihedral, resulting in a problematic or bad cosine series fit; thus, providing a more flexible and accurate fit by combining multiple dihedral conformational energies in a single dihedral, a strategy used in the original and modern OPLS force fields (Jorgensen et al., 1996; Lu et al., 2021). For example, a carboxylic acid with an alkyl tail has two dihedrals in the same rotation cycle; the C-C-C-O: (O: = oxygen without hydrogen) dihedral is set to zero while the C-C-O-H dihedral is fit (Jorgensen et al., 1996; Kamath et al., 2004; Lu et al., 2021). The `MoSDeF-dihedral-fit` (Crawford, Quach, et al., 2023) API fills the missing gap by providing a generalized and easy solution to fitting dihedrals in their commonly used forms and outputting the MM dihedral data points so users can fit other custom dihedral forms.

Acknowledgements

This research was partially supported by the National Science Foundation (grants OAC-1835713, OAC-1835874, and CBET 2052438). Atomfold LLC also donated research and development time and computational resources for this research and software. Wayne State University Grid provided some of the computational resources used in this work.

References

Allen, M. P., & Tildesley, D. J. (2017). *Computer simulation of liquids (2nd ed.)*. Oxford University Press.

Berthelot, D. (1898). Sur le mélange des gaz. *Comptes Rendus Hebd. Acad. Sci.*, *126*, 1703–1855. https://www.biodiversitylibrary.org/item/111847#page/757/mode/1up

Betz, R. (2016). *VMD-python*. Github. https://github.com/Eigenstate/vmd-python

Chen, S., Yi, S., Gao, W., Zuo, C., & Hu, Z. (2015). Force field development for organic molecules: Modifying dihedral and 1-*n* pair interaction parameters. *J Comput Chem.*, *36*, 376–384. https://doi.org/10.1002/jcc.23808

Crawford, B., Quach, C., Craven, N., Iacovella, C. R., McCabe, C., Cummings, P. T., & Potoff, J. (2023). *MoSDeF-dihedral-fit: A simple software package to fit dihedrals via the MoSDeF software*. Github. https://doi.org/10.5281/zenodo.14107384

Crawford, B., Timalsina, U., Quach, C. D., Craven, N. C., Gilmer, J. B., McCabe, C., Cummings, P. T., & Potoff, J. J. (2023). MoSDeF-GOMC: Python software for the creation of scientific workflows for the Monte Carlo simulation engine GOMC. *Journal of Chemical Information and Modeling*, *63*(4), 1218–1228. https://doi.org/10.1021/acs.jcim.2c01498

Crawford, B., Timalsina, U., Quach, C. D., Craven, N., Gilmer, J., Cummings, P. T., & Potoff, J. (2022). *MoSDeF-GOMC: Python software for the creation of scientific workflows for the Monte Carlo simulation engine GOMC*. Github. https://doi.org/10.5281/zenodo.14266177

Cummings, P. T., McCabe, C., Iacovella, C. R., Ledeczi, A., Jankowski, E., Jayaraman, A., Palmer, J. C., Maginn, E. J., Glotzer, S. C., Anderson, J. A., Siepmann, J. I., Potoff, J., Matsumoto, R. A., Gilmer, J. B., DeFever, R. S., Singh, R., & Crawford, B. (2021). Open-source molecular modeling software in chemical engineering, with focus on the Molecular Simulation Design Framework (MoSDeF). *AICHE J.*, *67(3)*, e17206. https://doi.org/10.1002/aic.17206

Errington, J. R., & Panagiotopoulos, A. Z. (1999). A new intermolecular potential model for the *n*-alkane homologous series. *The Journal of Physical Chemistry B*, *103*(30), 6314–6322. https://doi.org/10.1021/jp990988n

Frenkel, D., & Smit, B. (2002). *Understanding molecular simulation from algorithms to applications (2nd ed.)*. Academic Press.

Friederich, P., Konrad, M., Strunk, T., & Wenzel, W. (2018). Machine learning of correlated dihedral potentials for atomistic molecular force fields. *Scientific Reports*, *8*, 2559. https://doi.org/10.1038/s41598-018-21070-0

Frisch, M. J., Trucks, G. W., Schlegel, H. B., Scuseria, G. E., Robb, M. A., Cheeseman, J. R., Scalmani, G., Barone, V., Petersson, G. A., Nakatsuji, H., Li, X., Caricato, M., Marenich, A. V., Bloino, J., Janesko, B. G., Gomperts, R., Mennucci, B., Hratchian, H. P., Ortiz, J. V., … Fox, D. J. (2016). *Gaussian 16 Revision C.01*.

GMSO: General molecular simulation object. (2019). Github. https://doi.org/10.5281/zenodo.12533861

Good, R. J., & Hope, C. J. (1970). New combining rule for intermolecular distances in intermolecular potential functions. *J. Chem. Phys.*, *53*, 540–543. https://doi.org/10.1063/1.1674022

Guvench, O., & MacKerell, A. D. (1998). Automated conformational energy fitting for force-field development. *J. Mol. Model.*, *14*, 667–679. https://doi.org/10.1007/s00894-008-0305-0

Hemmen, A., & Gross, J. (2015). Transferable anisotropic united-atom force field based on the mie potential for phase equilibrium calculations: *N*-alkanes and *n*-olefins. *The Journal of Physical Chemistry B*, *119*(35), 11695–11707. https://doi.org/10.1021/acs.jpcb.5b01354

Hirst, J. D., Glowacki, D. R., & Baaden, M. (2014). Molecular simulations and visualization: Introduction and overview. *Faraday Discussions*, *169*, 9–22. https://doi.org/10.1039/c4fd90024c

Hollingsworth, S. A., & Dror, R. O. (2018). Molecular dynamics simulation for all. *Neuron*, *99*, 1129–1143. https://doi.org/10.1016/j.neuron.2018.08.011

Horton, J. T., Boothroyd, S., Wagner, J., Mitchell, J. A., Gokey, T., Dotson, D. L., Kumar, P., Ramaswamy, B. K., M., M., Chodera, J. D., Anwar, J., Mobley, D. L., & Cole, D. J. (2022). Open Force Field BespokeFit: Automating Bespoke Torsion Parametrization at Scale. *J. Chem. Inf. Mod.*, *62*, 5622–5633. https://doi.org/10.1021/acs.jcim.2c01153

Huang, L., & Roux, B. (2013). Automated force field parameterization for nonpolarizable and polarizable atomic models based on ab initio target data. *Journal of Chemical Theory and Computation*, *9*(8), 3543–3556. https://doi.org/10.1021/ct4003477

Humphrey, W., Dalke, A., & Schulten, K. (1996). VMD – Visual Molecular Dynamics. *Journal of Molecular Graphics*, *14*, 33–38. https://doi.org/10.1016/0263-7855(96)00018-5

Jorgensen, W. L., Maxwell, D. S., & Tirado-Rives, J. (1996). Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *J. Amer. Chem. Soc.*, *118(45)*, 11225–11236. https://doi.org/10.1021/ja9621760

Kamath, G., Cao, F., & Potoff, J. J. (2004). An improved force field for the prediction of the vapor−liquid equilibria for carboxylic acids. *The Journal of Physical Chemistry B*, *108*(37), 14130–14136. https://doi.org/10.1021/jp048581s

Kania, A., Sarapata, K., Gucwa, M., & Wójcik-Augustyn, A. (2021). Optimal solution to the torsional coefficient fitting problem in force field parametrization. *The Journal of Physical Chemistry A*, *125*(12), 2673–2681. https://doi.org/10.1021/acs.jpca.0c10845

Koneru, B., Swapnalin, J., Banerjee, P., Naidu, K. C. B., & Kumar, N. S. (2022). Materials under extreme pressure: Combining theoretical and experimental techniques. *The European Physical Journal Special Topics*, *231*, 4221. https://doi.org/10.1140/epjs/s11734-022-00569-8

Kumar, G., Mishra, R. R., & Verma, A. (2022). Introduction to molecular dynamics simulations. In A. Verma, S. Mavinkere Rangappa, S. Ogata, & S. Siengchin (Eds.), *Forcefields for atomistic-scale simulations: Materials and applications* (pp. 1–19). Springer Nature Singapore. https://doi.org/10.1007/978-981-19-3092-8_1

Lorentz, H. A. (1881). Ueber die Anwendung des Satzes vom Virial in der kinetischen Theorie der Gase. *Ann. D. Phys.*, *12*, 127–136. https://doi.org/10.1002/andp.18812480110

Louie, S. G., Chan, Y.-H., Jornada, F. H. da, Li, Z., & Qiu, D. Y. (2021). Discovering and understanding materials through computation. *Nature Materials*, *20*, 728. https://doi.org/10.1038/s41563-021-01015-1

Lu, C., Wu, C., Ghoreishi, D., Chen, W., Wang, L., Damm, W., Ross, G. A., Dahlgren, M. K., Russell, E., Von Bargen, C. D., Abel, R., Friesner, R. A., & Harder, E. D. (2021). OPLS4: Improving force field accuracy on challenging regimes of chemical space. *Journal of Chemical Theory and Computation*, *17*(7), 4291–4300. https://doi.org/10.1021/acs.jctc.1c00302

Martin, M. G., & Siepmann, J. I. (1998). Transferable potentials for phase equilibria. 1. United-atom description of *n*-alkanes. *J. Phys. Chem. B*, *102*(14), 2569–2577. https://doi.org/10.1021/jp972543+

Mayne, C. G. M., Saam, J, Schulten, K., Tajkhorshid, E., & Gumbart, J. C. (2013). Rapid parameterization of small molecules using the force field toolkit. *J. Comp. Chem.*, *34*, 2757–2770. https://doi.org/10.1002/jcc.23422

Mielke, R. R., Leathrum, J. F., Collins, A. J., & Audette, M. A. (2019). Overview of computational modeling and simulation. In D. Nestel, J. Hui, K. Kunkler, M. W. Scerbo, & A. W. Calhoun (Eds.), *Healthcare simulation research: A practical guide* (pp. 39–47). Springer International Publishing. https://doi.org/10.1007/978-3-030-26837-4_6

Nejahi, Y., Soroush Barhaghi, M., Mick, J., Jackman, B., Rushaidat, K., Li, Y., Schwiebert, L., & Potoff, J. (2019). GOMC: GPU Optimized Monte Carlo for the simulation of phase equilibria and physical properties of complex fluids. *SoftwareX*, *9*, 20–27. https://doi.org/10.1016/j.softx.2018.11.005

Nejahi, Y., Soroush Barhaghi, M., Schwing, G., Schwiebert, L., & Potoff, J. (2021). Update 2.70 to "GOMC: GPU Optimized Monte Carlo for the simulation of phase equilibria and physical properties of complex fluids". *SoftwareX*, *13*, 100627. https://doi.org/10.1016/j.

softx.2020.100627

Potoff, J. J., & Bernard-Brunel, D. A. (2009). Mie potentials for phase equilibria calculations: Application to alkanes and perfluoroalkanes. *The Journal of Physical Chemistry B*, *113*(44), 14725–14731. https://doi.org/10.1021/jp9072137

Schmid, N., Eichenberger, A. P., Choutko, A., Riniker, S., Winger, M., Mark, A. E., & Gunsteren, W. F. van. (2011). Definition and testing of the GROMOS force-field versions 54A7 and 54B7. *European Biophysics Journal*, *40*, 843. https://doi.org/10.1007/s00249-011-0700-9

Siegfried, R. (2014). Introduction. In *Modeling and simulation of complex systems: A framework for efficient agent-based modeling and simulation* (pp. 1–8). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-07529-3_1

Stone, J., Gullingsrud, J., Grayson, P., & Schulten, K. (2001). A system for interactive molecular dynamics simulation. In J. F. Hughes & C. H. Séquin (Eds.), *2001 ACM symposium on interactive 3D graphics* (pp. 191–194). ACM SIGGRAPH. https://doi.org/10.1145/364338.364398

Summers, A. Z., Gilmer, J. B., Iacovella, C. R., Cummings, P. T., & Mccabe, C. (2020). MoSDeF, a Python Framework Enabling Large-Scale Computational Screening of Soft Matter: Application to Chemistry-Property Relationships in Lubricating Monolayer Films. *J. Chem. Theor. Comput.*, *16*(3), 1779–1793. https://doi.org/10.1021/acs.jctc.9b01183

Swai, R. E. (2020). A review of molecular dynamics simulations in the designing of effective shale inhibitors: Application for drilling with water-based drilling fluids. *Journal of Petroleum Exploration and Production Technology*, *10*, 3515. https://doi.org/10.1007/s13202-020-01003-2

Timalsina, U. (2022). *Forcefield-utilities*. Github. https://doi.org/10.5281/zenodo.10494867

Vanommeslaeghe, Kenno, Guvench, O., & Alexander D. MacKerell, Jr. (2014). Molecular mechanics. *Current Pharmaceutical Design*, *20*, 3281–3292. https://doi.org/10.2174/13816128113199990600

Vanommeslaeghe, K., Hatcher, E., Acharya, C., Kundu, S., Zhong, S., Shim, J., Darian, E., Guvench, O., Lopes, P., Vorobyov, I., & A. D. MacKerell, Jr. (2010). CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *J Comput Chem.*, *31*, 671–690. https://doi.org/10.1002/jcc.21367

Vermeyen, T., Cunha, A., Bultinck, P., & Herrebout, W. (2023). Impact of conformation and intramolecular interactions on vibrational circular dichroism spectra identified with machine learning. *Communications Chemistry*, *6*, 148. https://doi.org/10.1038/s42004-023-00944-z

Weiner, Scott J., Kollman, P. A., Case, D. A., Singh, U. C., Ghio, C., Alagona, G., Profeta, S., & Weiner, P. (1984). A new force field for molecular mechanical simulation of nucleic acids and proteins. *Journal of the American Chemical Society*, *106*(3), 765–784. https://doi.org/10.1021/ja00315a051

Weiner, S. J., Kollman, P. A., Nguyen, N. D. T., & Case, D. A. (1986). An all atom force field for simulations of proteins and nucleic acids. *J. Comp. Chem.*, *7*, 230–252. https://doi.org/10.1002/jcc.540070216

Yu, D., & Pahl, E. (2023). Melting of atomic materials under high pressures using computer simulations. *Advances in Physics: X*, *8*(1), 2235060. https://doi.org/10.1080/23746149.2023.2235060