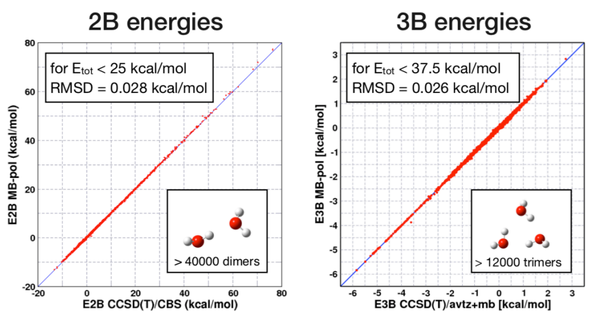
SI2-SSE: Enabling Chemical Accuracy in Computer Simulations: An Integrated Software Platform for Many-Body Molecular Dynamics

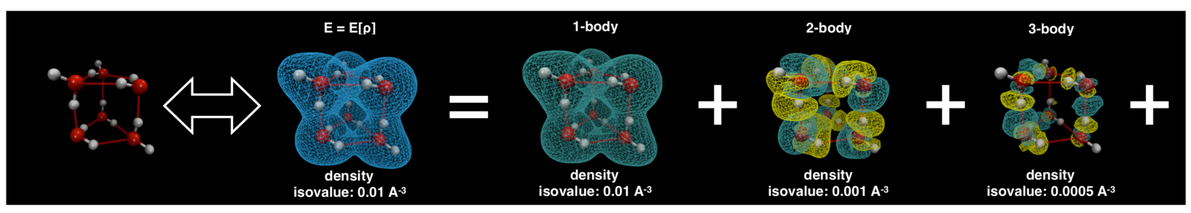
Award #: 1642336
PI: Francesco Paesani, Co-PIs: Andreas W. Götz, Andrea Zonca
Institution: University of California, San Diego

Many-body molecular dynamics: Chemical Accuracy Across Different Phases

$$E_N = \sum_i^N V^{1B}(i) \quad \text{1-body}$$
$$+ \sum_{i<j}^N V^{2B}(i,j) \quad \text{2-body}$$
$$+ \sum_{i<j<k}^N V^{3B}(i,j,k) \quad \text{3-body}$$
$$+ \ldots$$
$$+ V^{NB}(1,\ldots,N) \quad \text{N-body}$$

explicit 1B term
+
N-body induction
+
explicit 2B and 3B terms from CCSD(T)/CBS





Key features

- No recourse to empirical parameters
- Potential energy surface entirely derived from highly accurate correlated electronic structure data
- Enables predictive molecular simulations from gas phase to the condensed phase

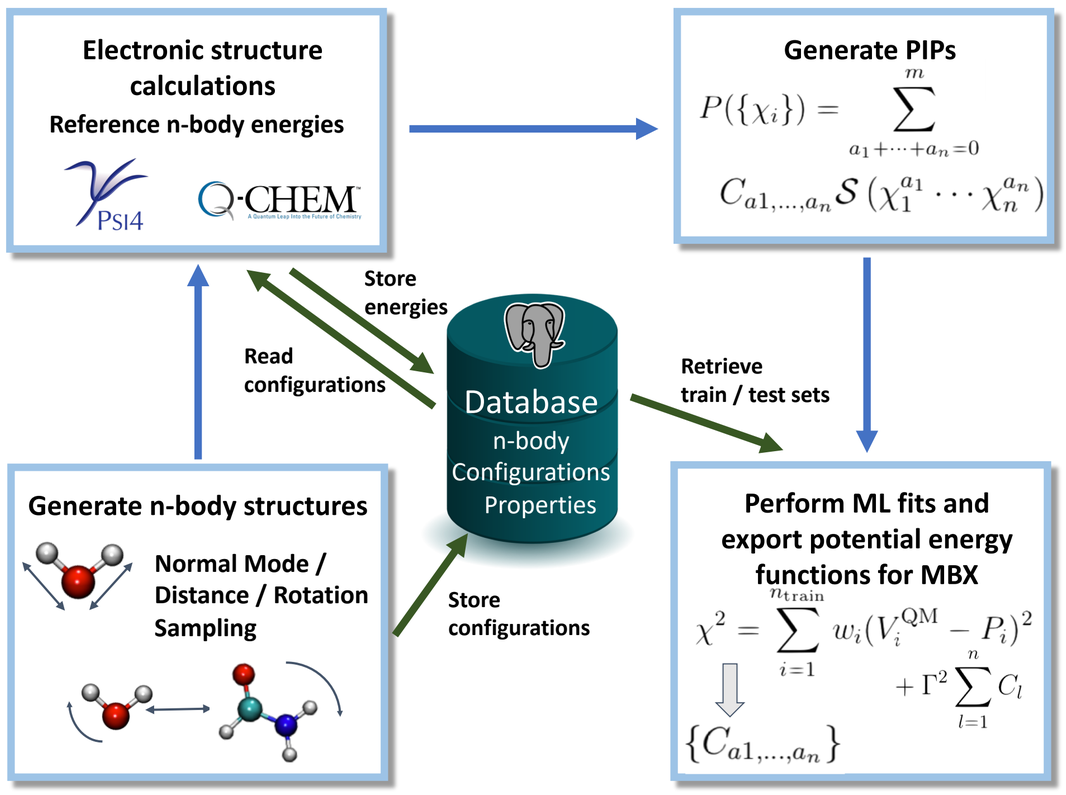
Goal of this project

- Make MB-MD methodology widely available to a broad scientific user community

MB-MD in OpenMM:
http://paesanigroup.ucsd.edu/software/mbpol_openmm.html

- *J. Phys. Chem. Lett.* 3, 3765 (2012)
- *J. Chem. Theory Comput.* 9, 1103 (2013)
- *J. Chem. Theory Comput.* 9, 4844 (2013)
- *J. Chem. Theory Comput.* 9, 5395 (2013)
- *J. Chem. Theory Comput.* 10, 1599 (2014)
- *J. Chem. Theory Comput.* 10, 2906 (2014)
- *J. Chem. Theory Comput.* 11, 1145 (2015)
- *J. Chem. Phys.* 145, 194504 (2016)

MB-fit: Workflow system for machine learning of many-body potential energy functions



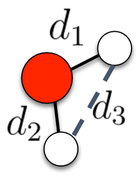
$$P(\{\chi_i\}) = \sum_{a_1+\cdots+a_n=0}^{m} C_{a1,\ldots,a_n} \mathcal{S}\left(\chi_1^{a_1}\cdots\chi_n^{a_n}\right)$$

$$\chi^2 = \sum_{i=1}^{n_{\text{train}}} w_i (V_i^{\text{QM}} - P_i)^2 + \Gamma^2 \sum_{l=1}^{n} C_l$$

$$\{C_{a1,\ldots,a_n}\}$$

Features

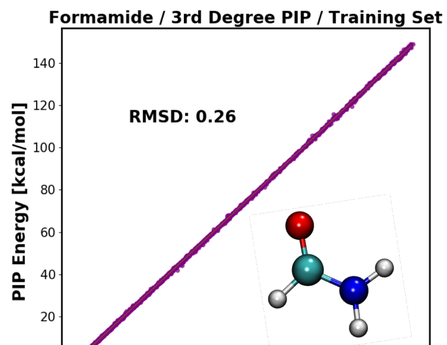
- Python based workflow system
- Jupyter notebooks for user documentation
- Centralized data storage via PostgreSQL
- Reference n-body structure generators
- Drivers for electronic structure calculations
- Efficient generator for permutationally invariant polynomials (PIPs)

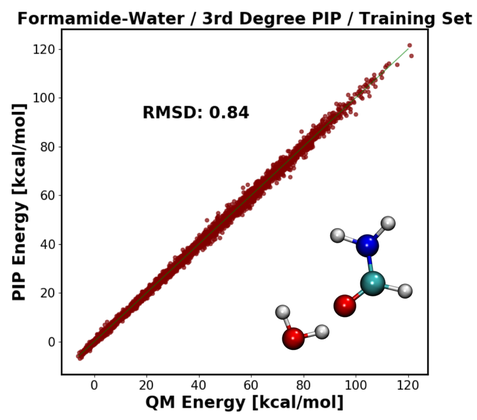


Permutationally invariant polynomials (PIPs)

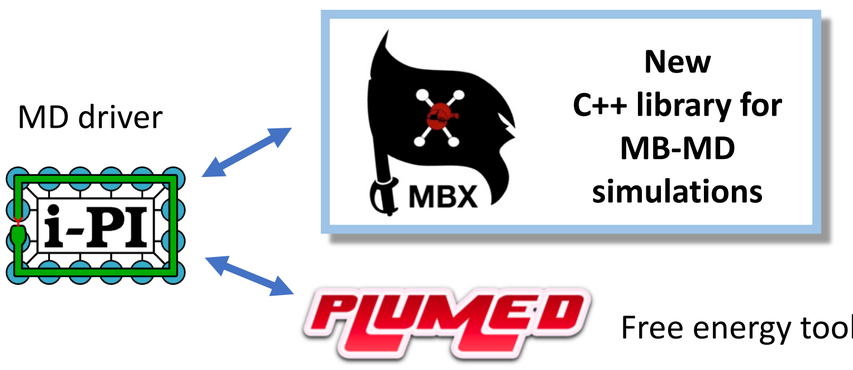
$$\chi_i = \exp(-k_i d_i) \quad \text{or} \quad \chi_i = \exp(-k_i d_i)/d_i$$

$$P(\{\chi_i\}) = \sum_{a_1+\cdots+a_n=0}^{m} C_{a1,\ldots,a_n} \mathcal{S}\left(\chi_1^{a_1}\cdots\chi_n^{a_n}\right)$$





MBX: High-performance C++ / OpenMP code for MB-MD simulations



OpenMP Enabling HPC since 1997

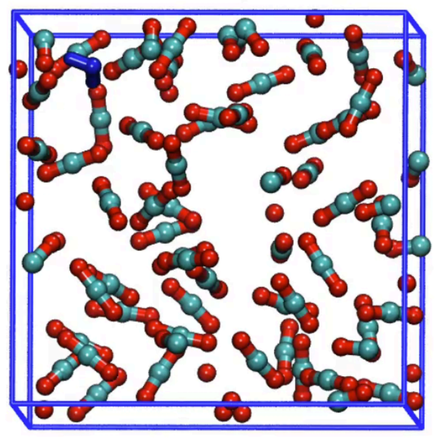
- OpenMP parallel for multi-core CPUs
- MPI support for multi-node clusters in progress

Features

- Modern C++ implementation
- OpenMP parallel
- Periodic boundary conditions
- Long-range electrostatics and dispersion energies via helPME library (https://github.com/andysim/helpme)
- User-friendly API for integration with simulation codes
- MD driver via i-PI
- Free energies via PLUMED

Software available at
http://paesanigroup.ucsd.edu/software/mbx.html

Application example:
MBX powered condensed phase simulation of dilute H_2O in CO_2 with coupled cluster accuracy (chemrxiv.11288465.v1)



Broader Impacts

- Training of a large number of postdoctoral scholars, graduate students, undergraduate students and high-school interns in computational research and software engineering
- Open source software tools for predictive molecular simulations for the broader scientific community

Acknowledgments

Marc Riera, Ethan Bull-Vulpe, Andy Simmonett, Daniel Smith

MolSSI
SDSC SAN DIEGO SUPERCOMPUTER CENTER
XSEDE Extreme Science and Engineering Discovery Environment

ACI-1642336