David Graff

Brooklyn, NY | davidegraff62@gmail.com | github.com/davidegraff

Education

Ph.D. Theoretical Chemistry — May 2023
Harvard University
Advisors: Connor Coley (MIT), Eugene Shakhnovich (Harvard)
Thesis: "Accelerating Exploration in Virtual Chemical Libraries"

A.B. Chemistry, *magna cum laude* — Jun 2018
Princeton University
Advisor: Robert Knowles
Thesis: "Development of Two Novel Reactions for Alkene Hydroamination Enabled by Proton-Coupled Electron Transfer"
Certificate: Applications of Computing

Research Experience

Machine Learning Research Scientist II — Sep 2023—Present
Atomwise, San Francisco, CA

Postdoctoral Associate — Jun 2023—Aug 2023
Advisor: Connor Coley
MIT Department of Chemical Engineering, Cambridge MA

Graduate Research Assistant — Jan 2020—May 2023
Advisors: Connor Coley (MIT), Eugene Shakhnovich (Harvard)
MIT Department of Chemical Engineering / Harvard Department of Chemistry and Chemical Biology, Cambridge, MA

Machine Learning Research Intern — Jun 2022—Aug 2022
Roivant Discovery, Boston, MA

Undergraduate Research Assistant — Feb 2017—Apr 2018
Advisors: Robert Knowles | Collaborators: Casey Roos, Qilei Zhu
Princeton University Department of Chemistry, Princeton, NJ

Harvard University BCMP Summer Scholar — Jun 2016—Aug 2016
Advisor: Eric Fischer | Collaborator: Radek Nowak
Harvard Medical School / Dana Farber Cancer Institute, Boston, MA

Publications

1. Dicks, L.; Graff, D.E.; Jordan, K.E.; Coley C.W.; Pyzer-Knapp, E.O. A physics-inspired approach to the understanding of molecular representations and models. *Mol. Syst. Des. Eng.* **2024**, *9*, 449
2. J.C. Fromer; D.E. Graff; C.W. Coley. Pareto optimization to accelerate multi-objective virtual screening. *Digi. Discov.* **2024**, *3*, 467
3. Heid, E.; Greenman, K. P.; Chung, Y.; Li, S.-C; Graff, D.E.; Vermeire, F.H.; Wu, H.; Green, W.H.; McGill C.J. Chemprop: A Machine Learning Package for Chemical Property Prediction. *J. Chem. Inf. Model.* **2024**, *64*, 1, 9–17
4. Graff, D.E.; Pyzer-Knapp, E.O.; Shakhnovich, E.I.; Coley, C.W. Evaluating the Roughness of Structure-Property Relationships using Pretrained Molecular Representations. *Digi. Discov.* **2023**, *2* (5), 1452
5. Aldeghi, M.; Graff, D.E.; Frey, N.C.; Morrone, J.A.; Pyzer-Knapp, E.O.; Jordan, K.E.; Coley, C.W. Roughness of Molecular Property Landscapes and Its Impact on Modellability. *J. Chem. Inf. Model.* **2022**, *62* (19), 4660
6. Graff, D.E.; Aldeghi, M.; Morrone, J.A.; Jordan, K.E.; Pyzer-Knapp, E.O.; Coley, C.W. Self-Focusing Virtual Screening with Active Design Space Pruning. *J. Chem. Inf. Model.* **2022**, *62* (16), 3854
   ACS Editors' Choice
7. Graff, D.E.; Coley, C.W. pyscreener: A Python Wrapper for Computational Docking Software. *J. Open Source Softw.* **2022**, *7* (71), 3950
8. Graff, D.E.; Shakhnovich, E.I., Coley, C.W. Accelerating High-Throughput Virtual Screening Through Molecular Pool-Based Active Learning. *Chem. Sci.* **2021**, *12*, 7866
   Featured in State of AI Report 2021
9. Roos, C.B.; Demaerel, J.; Graff, D.E.; Knowles, R.R. Enantioselective Hydroamination of Alkenes with Sulfonamides Enabled by Proton-Coupled Electron Transfer. *J. Am. Chem. Soc.* **2020**, *142* (13), 5974
10. Zhu, Q.; Graff, D.E.; Knowles, R.R. Intermolecular Anti-Markovnikov Hydroamination of Unactivated Alkenes with Sulfonamides Enabled by Proton-Coupled Electron Transfer. *J. Am. Chem. Soc.* **2018**, *140* (2), 741.

Presentations

Oral

1. "MolPAL: Software for Sample Efficient High-Throughput Virtual Screening" — Dec 2022
   NeurIPS AI4Mat, *New Orleans, LA*
   Spotlight Talk
2. "Chemprop: New features and updates" — Nov 2022
   MIT MLPDS Consortium Meeting, *Boston, MA*,
3. "Open-source software tools to accelerate docking-based molecular discovery" — Aug 2022
   Multidisciplinary and Data-Driven Software Development, ACS National Meeting, *Chicago, IL*
4. "Accelerating HTVS through molecular pool-based active learning" — Apr 2021
   The Future of AI in Chemistry, ACS National Meeting, *Virtual*

5. "Apo crystal structure of DDB1ΔB" Aug 2016
   Harvard Medical School BSSP Research Symposium, *Boston, MA*

Poster

1. "The questionable advantages of pretrained chemical models in QSPR" Mar 2023
   OpenEye CUP XXII, *Santa Fe, NM*
2. " Self-focusing virtual screening with active design space pruning" May 2022
   MIT-IBM Watson AI Lab open house, *Boston, MA*
3. "A Tale of Two Sulfonamides. Princeton Chemistry Summer Research Symposium Aug 2017
   Princeton, NJ, August 2017

Teaching

Harvard University, *Cambridge, MA*

Teaching Fellow, Foundational Chemistry and Biology (LPS A) Fall 2020
Teaching Fellow, Experimental Inorganic Chemistry (CHEM 145) Fall 2019
Teaching Fellow, Biochemistry (CHEM 27) Spring 2019

Princeton University, *Princeton, NJ*

McGraw Center Head Tutor, Organic Chemistry (CHM 303/304/304B) Fall 2017–Spring 2018
Undergraduate Teaching Assistant, Organic Chemistry II (CHM 304) Spring 2018
Undergraduate Teaching Assistant, Organic Chemistry II (CHM 304B) Spring 2017, Spring 2018
Undergraduate Teaching Assistant, Organic Chemistry I (CHM 303) Fall 2016, Fall 2017
Peer Tutor, Algorithms and Data Structures (COS 226) Spring 2018
Peer Tutor, Introduction to Computer Science (COS 126) Fall 2016–Spring 2018

Skills

Technical

Machine Learning: graph neural networks, active learning, optimization, uncertainty quantification, generative models, large language models, pretrained, self-supervised learning
Chemistry: QSPR modelling, molecular docking, structure-based drug design

Programming

Languages: Python, Bash, C, Java, OCaml
Frameworks: PyTorch, Tensorflow, RDKit, Ray, Dask, NumPy, SciPy, Scikit-Learn, Pandas, Jupyter, Matplotlib, Seaborn
Software: SLURM, Docker, AutoDock Vina, DOCK6, PyMol
Practices: HPC, CI/CD, automated deployment, packaging, containerization, testing, version control, microservices