Multi-Fidelity Computer-Aided Molecular Design

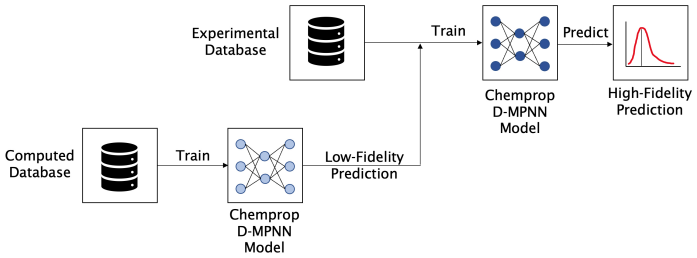
Kevin P. Greenman

Ph.D. Candidate, Department of Chemical Engineering, MIT

kpg@mit.edu

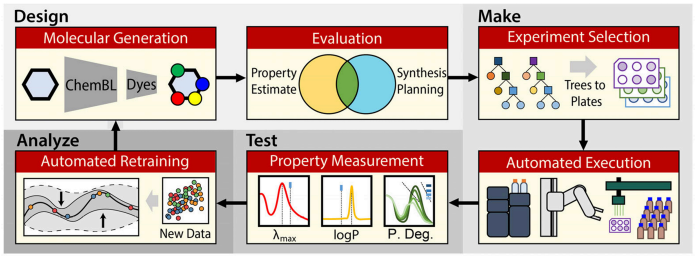
Prior Work

Multi-fidelity prediction of molecular optical peaks



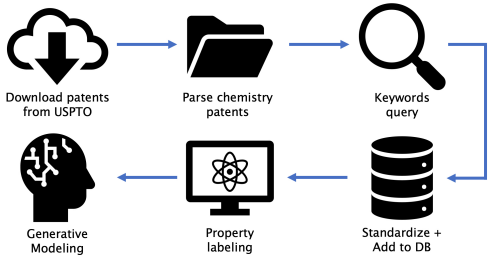
Greenman, K. P., Green, W. H., & Gómez-Bombarelli, R. (2022). Multi-fidelity prediction of molecular optical peaks with deep learning. *Chemical Science*, *13*(4), 1152-1162.

Automated closed-loop active learning



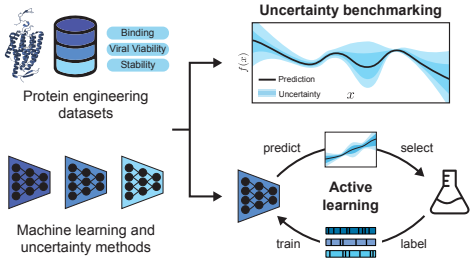
Koscher, B.*, Canty, R. B.*, McDonald, M. A.*, **Greenman, K. P.**, McGill, C. J., Bilodeau, C. L., ... & Jensen, K. F. (2023). Autonomous, multi-property-driven molecular discovery: from predictions to measurements and back. *ChemRxiv.*

Chemical patent data extraction and generative modeling



Subramanian, A.*, **Greenman, K. P.***, Gervaix, A., Yang, T., & Gómez-Bombarelli, R. (2023). Automated patent extraction powers generative modeling in focused chemical spaces. *Digital Discovery*.

Uncertainty in ML for protein engineering



Greenman, K. P., Amini, A. P., & Yang, K. K. (2023). Benchmarking uncertainty quantification for protein engineering. *bioRxiv*.

Chemprop software development

chemprop

Heid, E., **Greenman, K. P.,** Chung, Y., Li, S. C., Graff, D. E., Vermeire, F. H., ... & McGill, C. J. (2023). Chemprop: A machine learning package for chemical property prediction. *ChemRxiv*.

Future Research Vision

Create benchmark datasets for multi-fidelity, multi-objective, batch active learning

- **Progress often driven by benchmarks**
- **Enabling quantitative comparison between methods**

Design new molecular optical probes for biomedical imaging applications

- **Near-infrared absorption and emission**
- **Multi-objective design:** quantum yield, solubility, toxicity, etc.

Develop more robust tools for natural language explanations of chemistry

- **Explanations of uncertainty & active learning batch acquisition**
- **Suggestions of design choices**
- **Human-computer interaction**

Teaching

Curriculum Design

- **Introduction to Research:** undergraduate seminar + lab course (taught Winter 2019)

Teaching Experience

- **Machine Learning for Molecular Engineering:** Teaching assistant (graduate + undergraduate)
- **Fluid Mechanics:** Instructional aide (undergraduate)
- **MIT ChemE Teach-Off Competition:** 2021 first place awardee

Pedagogical Training

- **MIT Grad Teaching Development Tracks Certificate Program**
- **University of Michigan Science Learning Center**

Community Resources

- **Awesome ChemE Education GitHub List**

Acknowledgements

I was supported by the National Science Foundation Graduate Research Fellowship Program under Grant No. 1745302 and by the DARPA Accelerated Molecular Discovery (AMD) program under contract HR00111920025. I acknowledge the MIT Engaging cluster and MIT Lincoln Laboratory Supercloud cluster at the Massachusetts Green High Performance Computing Center (MGHPCC) for providing high-performance computing resources.