

RISHAL AGGARWAL

🐦 @AggarwalRishal | ✉ rishal.aggarwal@gmail.com | 🗣 Rishal Aggarwal | 🌐 RishalAggarwal

Research Interests: Machine learning methods for accelerated molecular screening, sampling, and design.

EDUCATION

CMU-Pitt Computational Biology

PhD Student

Pittsburgh, USA

Aug. 2022 – April 2027 (Anticipated)

International Institute of Information Technology

MS by Research in Bioinformatics

Hyderabad, India

Jan. 2021 – July 2022

Birla Institute of Technology and Sciences, Pilani

Bachelor of Pharmacy

Rajasthan, India

Aug. 2016 – May 2020

ACADEMIC AND RESEARCH ACHIEVEMENTS

- Authored multiple publications in top venues including acceptances at **NeurIPS**, **ICLR**-Blogposts, **ICML**-Genbio, **ICML**-SPIGM, **ICML**-CompBio, ACS JCI, Wiley, BMC Biology etc; recognized with **2400+ citations** and an **h-index of 9**. Google Scholar [link](#).
- **Contributions to Machine Learning:** Developed novel, scalable simulation-free methods for training energy-based models. Achieved amortized Boltzmann sampling using flow matching and EBMs, and conducted one of the first comprehensive evaluations of leading energy-trained Boltzmann samplers.
- **Contributions to AI in Chemistry:** Contributed to several projects significantly including the widely used molecular docking software **GNINA**. Co-developed two novel pharmacophore elucidation methods: PharmRL (reinforcement learning-based) and Pharmacoforge (diffusion/flow matching-based). **PharmRL** was recognized with a **Spotlight Poster** at the MLDD Conference, MIT - Broad Institute 2023.

SELECTED PUBLICATIONS

- **Rishal Aggarwal**, David R. Koes, Nicholas M. Boffi, and Eric Vanden-Eijnden. Reactive Flux Matching: Mechanism Discovery and Adaptive Sampling of Rare Events. *Submitted NeurIPS 2026*
- **Rishal Aggarwal**, Jacky Chen, Nicholas M. Boffi, and David R. Koes. BoltzNCE: Learning Likelihoods for Boltzmann Generation with Stochastic Interpolants and Noise Contrastive Estimation. *NeurIPS 2025, ICML - GenBio 2025*
- **Rishal Aggarwal**, Daniel Peñaherrera, Justin Shao, Minhyek Jeon, David R. Koes. Models trained with unnormalized density functions: A need for a course correction. *ICLR - Blogpost Track 2025*
- Anish Sevekari*, **Rishal Aggarwal***, David R. Koes, Maria Chikina. Accelerating NCE Convergence with Adaptive Normalizing Constant Computation. *ICML - Structured Probabilistic Inference & Generative Modeling 2024*
- **Rishal Aggarwal**, and David R. Koes. PharmRL: Pharmacophore elucidation with deep geometric reinforcement learning. *Springer BMC Biology, 2024*
- **Rishal Aggarwal**, Akash Gupta, and U. Deva Priyakumar. Apobind: a dataset of ligand unbound protein conformations for machine learning applications in de novo drug design. *ICML - Computational Biology 2021*
- **Rishal Aggarwal**, Akash Gupta, Vineeth Chelur, C. V. Jawahar, and U. Deva Priyakumar. DeepPocket: Ligand Binding Site Detection and Segmentation using 3D Convolutional Neural Networks. *Journal of Chemical Information and Modeling, 2021.*
- Viraj Bagal, **Rishal Aggarwal**, P. K. Vinod, and U. Deva Priyakumar. MolGPT: Molecular Generation Using a Transformer-Decoder Model. *Journal of Chemical Information and Modeling, 2021.*
- Andrew T McNutt, Paul Francoeur, **Rishal Aggarwal**, Tomohide Masuda, Rocco Meli, Matthew Ragoza, Jocelyn Sunseri, and David Ryan Koes. GNINA 1.0: molecular docking with deep learning. *Journal of cheminformatics, 13(1):1–20, 2021.*

WORK EXPERIENCE

Generative AI Research Intern

Achira.ai

May 2026 – August 2026

New York, NY

- **Energy-Based Steering of Co-Folding Models:** Implemented and evaluated multiple methods for steering generative co-folding models with energy functions, including Feynman–Kac steering, training-free guidance, Monte Carlo tree search, and flow-map-based guidance.
- **Energy-Based Fine-Tuning of Pretrained Generative Models:** Investigated methods for fine-tuning pretrained generative models with energy-based objectives to sample from Boltzmann distributions, including adjoint matching and non-equilibrium transport.

RESEARCH EXPERIENCE

Graduate RA / PhD student - Koes Group

CMU - Pitt Computational Biology

August 2022 – Present

Pittsburgh, USA

- **Amortized Boltzmann Sampling (BoltzNCE):** Developed a flow-matching and EBM-based method for efficient Boltzmann sampling of molecular systems. Achieved accurate estimation of free energy surfaces and energy histograms with orders-of-magnitude speedup over Boltzmann Generators and MD simulations. *Published at NeurIPS 2025, ICML-GenBio 2025.* [Link](#)
- **Reactive Flux Matching: Mechanism Discovery and Adaptive Sampling of Rare Events:** Proposed a novel framework that learns probability current ODEs from reactive path ensembles to identify the **optimal collective variable** governing a reaction. *Submitted to NeurIPS 2026.* [Link](#)
- **Boltzmann Sampling — Generative Models vs. MCMC:** Performed an in-depth evaluation of generative flow matching and diffusion models trained with energy functions, benchmarking them against MCMC approaches for Boltzmann sampling. Identified key insights on **transferability** and **domain adaptation**, outlining promising future research directions. *Published in ICLR Blogpost Track 2025.* [Link](#)
- **Accelerating Noise Contrastive Estimation (NCE):** Analyzed the impact of the NCE objective on the partition function parameter and developed a novel optimization routine leveraging **Bennett's Acceptance Ratio**, achieving significantly faster convergence. *Published at ICML — Structured Probabilistic Inference & Generative Modeling 2024.* [Link](#)
- **Pharmacophore Elucidation with Generative models:** Contributed significantly to the development of pharmacophore generation on protein binding sites using diffusion/flow matching. *Published Frontiers in Bioinformatics.* [Link](#)
- **Pharmacophore Elucidation with Reinforcement Learning:** Developed a framework for pharmacophore elucidation on protein binding pockets using geometric SO(3)-equivariant neural networks and Deep-Q Learning. *Published Spring BMC Biology 2024.* [Link](#)
- **Molecular Docking with Deep Learning (GNINA):** Contributed extensively to the development and validation of a widely adopted deep learning–based docking framework, now applied across numerous drug discovery projects. *Published in Journal of Cheminformatics, 2021 and 2024.* [Link](#)

Graduate Research Assistant - Deva Lab

Center for Computational Natural Sciences and Bioinformatics

May 2020 – August 2022

IIITH, India

- **Binding Pocket Identification (DeepPocket):** Trained Convolutional Neural Networks to detect and segment ligand binding sites on 3D protein structures, achieving state-of-the-art performance. *Published in ACS JCI.* [Link](#)
- **Molecular Generation with Transformers (MolGPT):** Demonstrated the effectiveness of attention-based models for molecular generation, enabling inverse design of molecules conditioned on user-defined scaffolds and property values. *Published in ACS JCI.* [Link](#)
- **APOBind — A Dataset of Apo–Holo Structures:** Created a benchmark dataset of unbound (apo) protein structures corresponding to protein–ligand complexes in PDBbind, enabling improved evaluation of binding affinity prediction and binding site detection methods under realistic apo conditions. Demonstrated that popular docking algorithms perform poorly on this dataset, underscoring its value for advancing method development. *Published in ICML -Computational Biology Workshop 2021.* [Link](#)
- **Hidden Biases in Binding Affinity Datasets:** Investigated systematic biases in protein–ligand datasets that influence the performance of ML models, providing a more holistic assessment of deep learning approaches for virtual screening. *Published in ACS Omega.* [Link](#)