Sihyun Park

sihyunpa@andrew.cmu.edu | sihyunpark99.github.io | (412)-513-2012

Research Interests

AI for Science, Generative AI, Computational Biology, Drug Discovery

Education

Carnegie Mellon University May 2027
M.S. in Automated Science *Pittsburgh, PA*

- Computational Biology Department, School of Computer Science
- Relevant coursework: 11-785 Introduction to Deep Learning; 10-701 Introduction to Machine Learning (PhD) Generative AI for Biomedicine, Modern Methods for Structure-Based Drug Discovery

Korea Advanced Institute of Science and Technology (KAIST) Feb. 2025
B.S. in Chemistry; Minor in Computer Science; Advanced Major in Chemistry *Daejeon, Korea*

- GPA: 4.01 / 4.30 (3.88 / 4.00), ***Summa Cum Laude***
- Military Service: Jan. 2020 – Jul. 2021

Georgia Institute of Technology Aug. 2022 – May 2023
Exchange Student *Atlanta, GA*

Publications

[1] **FragFlex: Fragment-Based Flexible Molecular Generation for Exploring Vast Chemical Space**
Sihyun Park, David Ryan Koes [link] *MoML 2026; NeurIPS AI4DD Workshop, 2026*

[2] **Discrete Diffusion Playground: A 2D Benchmark for Discrete Generative Models**
Justin Shao*, **Sihyun Park***, David Ryan Koes, Maria Chikina [link] *NeurIPS E&D Track, 2026*
*Equal contribution.

[3] **Towards Federated Learning Across Biobanks: Prototype Software from the 2026 Carnegie Mellon University–NVIDIA Hackathon**
James Mu, . . . , **Sihyun Park**, et al. [link] *BioHackrXiv, 2026*

Research Experience

Koes Lab, University of Pittsburgh Aug. 2025 – Present
MS Thesis Research (Advisor: David Ryan Koes) *Pittsburgh, PA*

- Developed **FragFlex**, a fragment-level graph diffusion model that dynamically adjusts molecular size while maintaining high chemical validity and diversity in large-scale generation (up to 10^7 samples).
- Developed the **Discrete Diffusion Playground**, a 2D benchmark that maps continuous distributions to binary sequences and back for direct visualization of generative model failures.

Isayev Lab, Carnegie Mellon University Sep. 2026 – Present
Research Project (Advisor: Olexandr Isayev) *Pittsburgh, PA*

- Working on the **AIMNet Docking** project, with a focus on molecular docking benchmark experiments.

Intelligent Chemistry Lab, KAIST Dec. 2023 – Jun. 2025
Undergraduate Research Program (Advisor: Woo Youn Kim) *Daejeon, Korea*

- Developed a transformer-based deep learning model to predict protein–ligand affinity and distance maps.
- Developed an active-learning-enhanced docking surrogate model for structure-based drug discovery.

Research Experience (continued)

NSF/NASA Center for Chemical Evolution, Georgia Tech Aug. 2022 – Jan. 2023
Undergraduate Research (Advisor: Nicholas V. Hud) *Atlanta, GA*

- Studied prebiotic chemistry using Solid-Phase Peptide Synthesis (SPPS), mass spectrometry, and HPLC.
- Synthesized proto-nucleobase-tagged tetrapeptides in research on candidate precursors relevant to the origin of RNA.

Center for Cell-Encapsulation Research, KAIST Jun. 2019 – Dec. 2019
Undergraduate Research (Advisor: Insung S. Choi) *Daejeon, Korea*

- Studied deep learning and developed a graph convolutional network (GCN) model based on three-dimensional molecular data.

Academic Service

Reviewer, ICML 2026 SPIGM Workshop 2026

Honors and Awards

The Best Book Club of KAIST (KAIST Presidential Award) 2024
Leadership Mileage Certificate (Diamond Level), KAIST 2024
The National Scholarship for Science and Engineering 2022 – 2023
Department Honors Scholarship, KAIST Fall 2021
Dean's List, KAIST Fall 2019

Skills

Programming Languages	Python, Java, C, Scala
Frameworks / Libraries	PyTorch, PyTorch Geometric, PyTorch Lightning, Keras, NumPy, Pandas
Tools / Software	Linux, PyMOL, RDKit, OpenBabel, Git, Adobe Photoshop, LaTeX
Chemical Analysis Equipment	HPLC, mass spectrometry, UV-VIS, NMR, PCR, SDS-PAGE
Language Proficiency	Fluent in English; Native in Korean