

SHUAIKE SHEN

Carnegie Mellon University, Pittsburgh, PA, United States, 15289
+1(412)-657-2355 | shuaikes@andrew.cmu.edu

EDUCATION

Carnegie Mellon University
Ph.D in Computational Biology

Pittsburgh, PA, U.S.
Aug. 2025 - Present

Zhejiang University
B.Eng in Computer Science and Technology (Honors)

Hangzhou, Zhejiang, China
Sept. 2021 - June. 2025

- Mixed Class, [Chu Kochen Honors College](#), GPA: 3.83/4.00

RESEARCH INTEREST

I am deeply interested in developing AI and machine learning methods for scientific discovery, particularly in biomedical applications. My research focuses on protein design and flow-matching models. I have also worked on topics related to scientific large language models, reinforcement learning, and crystal structure modeling.

RESEARCH EXPERIENCE

Research Assistant
Advisor: Prof. Olexandr Isayev & Prof. Lei Li

Carnegie Mellon University
Aug. 2025 - Present

- Working on low energy conformer generation and protein complex using diffusion models.

Research Intern
Advisor: Dr. Biqing Qi & Dr. Yuqiang Li

Shanghai AI Lab
May. 2025 - Aug. 2025

- Working on large language models for molecular spectra and developing molecular foundation models to bridge spectral data with molecular properties and structures.

Research Intern
Advisor: Prof. Ge Liu & Prof. Nicholas Ching Hai Wu

University of Illinois Urbana-Champaign
June. 2024 - Mar. 2025

- Working on latent flow matching model for protein sequence and structure co-design and developed online reinforcement algorithms for flow matching models.

Research Assistant
Advisor: Prof. Chunhua Shen & Prof. Hao Chen

Zhejiang University
Oct. 2022 - May. 2024

- Working on protein engineering and crystal structure modeling. I developed vector field protein encoder, multi-motif scaffolding method, RNA inverse folding paradigm, binding energy prediction model and crystal denoising pretraining framework.

PUBLICATIONS

* denotes equal contribution and † denotes corresponding author.

MolSpectLLM: A Molecular Foundation Model Bridging Spectroscopy, Molecule Elucidation and 3D Structure Generation Sep. 2025

- Shuaike Shen, Jiaqing Xie, Zhuo Yang, Antong Zhang, Shuzhou Sun, Ben Gao, Tianfan Fu†, Biqing Qi†, Yuqiang Li†
- Under review, [Paper](#), [Code](#)

SpectrumWorld: Artificial Intelligence Foundation for Spectroscopy Aug. 2025

- Zhuo Yang*, Jiaqing Xie*, Shuaike Shen, Daolang Wang, Yeyun Chen, Ben Gao, Shuzhou Sun, Biqing Qi, Dongzhan Zhou, Lei Bai, Linjiang Chen, Shufei Zhang, Jun Jiang†, Tianfan Fu†, Yuqiang Li†
- Under review, [Paper](#)

Intern-S1: A Scientific Multimodal Foundation Model Aug. 2025

- Intern-S1 team
- Technical Report, [Paper](#), [Code](#), [Huggingface](#)

From Sentences to Sequences: Rethinking Languages in Biological System • Ke Liu[*], Shuaike Shen[*], Hao Chen[†] • Paper, Code	July. 2025
Online Reward-Weighted Fine-Tuning of Flow Matching with Wasserstein Regularization • Jiajun Fan, Shuaike Shen, Chaoran Cheng, Yuxin Chen, Chumeng Liang, Ge Liu[†] • ICLR 2025, Paper	Jan. 2025
A Denoising Pre-training Framework for Accelerating Novel Material Discovery • Shuaike Shen, Ke Liu, Hao Chen[†] • AAAI 2025 AISI, Paper	Dec. 2024
Training Text-to-Image Flow Models on Self-Generated Data • Jiajun Fan, Shuaike Shen, Chaoran Cheng, Chumeng Liang, Ge Liu[†] • Under review	Nov. 2024
Physics-Informed Neural Networks for Unsupervised Binding Energy Prediction • Ke Liu[*], Shuaike Shen[*], Hao Chen[†] • Under review, Manuscript	Aug. 2024
Boost Your Crystal Model with Denoising Pre-training • Shuaike Shen[*], Ke Liu[*], Muzhi Zhu, Hao Chen[†] • ICML 2024 AI for Science Workshop, Paper, Project Page	June. 2024
FADiff: Floating Anchor Diffusion Model for Multi-motif Scaffolding • Ke Liu[*], Shuaike Shen[*], Weian Mao[*], Xiaoran Jiao, Zheng Sun, Hao Chen[†], Chunhua Shen[†] • ICML 2024, Paper, Project Page	May. 2024
De novo protein design using geometric vector field networks • Weian Mao[*], Muzhi Zhu[*], Zheng Sun[*], Shuaike Shen, Lin Yuanbo Wu, Hao Chen[†], Chunhua Shen[†] • ICLR 2024 spotlight, Paper, Project Page	Oct. 2023

SELECTED PROJECTS

VITON: High resolution virtual try on based on generative models <i>Scientific research projects of Chu Kochen Honors College advised by Prof. Hao Chen</i>	2023
MiniSQL: A database system implemented in C++ <i>The course project of Database systems and concepts</i>	2023
Parallel optimization of big data pivots selection based on vectorization and OpenMP <i>High Performance Computing 101 summer course ACM-China IPCC 2022</i>	2022

SKILLS

Programming Skills: Python, C/C++, PyTorch, SQL

Language Skills: Mandarin (Native), English (TOEFL 104)