JINGYI LI

Email: jingyili.mechanics@gmail.com
Mobile: +1-(814)321-2568
LinkedIn: linkedin.com/in/Jingyi-Victor-Li

Qualification Summary

- Ph.D. in Computational Biophysics specializing in **AI/ML for structural biology** and **molecular dynamics simulations**.
- Build **geometric diffusion models** and **Graph Neural Networks (EGNN/GCN)** for RNA structure prediction and refinement.
- Construct large-scale structural databases with **graph-based search algorithms** over hundreds of thousands of motifs.
- Proficient in **Python, PyTorch, LAMMPS**, and **HPC systems**; experienced in applying deep learning to biological and mechanical systems.

Education

- **Penn State University** — State College, PA
  Ph.D., Engineering Science and Mechanics — *Jun 2022 – May 2026*
- **Dalian University of Technology** — Liaoning, China
  MPhil, Engineering Mechanics — *Sep 2019 – Jun 2022*
- **Dalian University of Technology** — Liaoning, China
  Bachelor, Engineering Mechanics — *Sep 2015 – Jun 2019*

Technical Skills

- **Machine Learning / Deep Learning:** PyTorch, Diffusion Models (DDPM), Graph Neural Networks (EGNN, GCN), CNN
- **Programming:** Python, C++, MATLAB, SQL, Bash/Shell scripting, Git
- **Simulation & Modeling:** LAMMPS, GROMACS (Molecular Dynamics), Abaqus (FEM), MolProbity, PyMOL
- **Infrastructure & Tools:** Linux, HPC clusters (SLURM/Roar), LLM-assisted development (Claude Code, custom MCP servers, skills, and agents for research automation)
- **Bioinformatics:** PDB data processing, RNA structure analysis

Research Experience

- **Pennsylvania State University** — State College, PA
  Graduate Research Assistant, Advisors: Prof. Nikolay V. Dokholyan and Prof. Sulin Zhang — *Jun 2022 – May 2026*

ChironRNA: RNA structure refinement via E(3)-equivariant diffusion (First Author)

- Designed and implemented a **DDPM-based diffusion framework** with **E(3)-equivariant Graph Neural Networks (EGNN)** in **Python/PyTorch** for resolving steric clashes and reconstructing missing atoms in RNA structures.
- Developed a **hierarchical architecture** combining all-atom and coarse-grained (five-point nucleotide representation) diffusion models, where the coarse-grained model recovers 23.7% of cases where the all-atom pipeline stalls.
- Implemented conditional generation pipeline using **MolProbity** diagnostics to identify clash regions and selectively regenerate distorted substructures while preserving valid regions as constraints.
- Achieved **80% clash reduction** on over 80% of the test set and **100% atom reconstruction rate** for structures under 200 nucleotides. Benchmarked on 90 RNA motifs with low RMSD and chemically accurate coordinates.

ATLAS: Large-scale 3D RNA motif library with graph search (First Author)

- Built an automated data pipeline to process **8,791 RNA structures** from PDB, extracting and classifying **433,996 motifs** across five categories (hairpin loops, bulges, internal loops, junctions, pseudoknots).
- Engineered three representations per motif: 3D atomic coordinates, nucleotide-level graphs with non-Watson-Crick interactions, and graphs without non-WC interactions. Implemented **graph isomorphism search algorithm** for querying motifs by user-defined substructures.
- Revealed that non-WC interactions significantly increase structural diversity, e.g., tetraloop hairpins alone exhibit **1,487 unique conformations**, most of which were invisible to prior libraries.
- Developed web interface for community access to search, visualize, and download RNA motifs, supporting both canonical and non-canonical interaction patterns.

Molecular dynamics of plant epidermal cell wall mechanics

- Performed large-scale **MD simulations** in **LAMMPS** to investigate the mechanical behavior of cellulose microfibers in plant epidermal cell walls, building upon the methodology of Zhang et al. (*Science*, 2021).
- Modeled shear stress-sliding relationships in cellulose bundles, capturing the telescopic sliding mechanism that enables cell walls to maintain both strength and extensibility during plant growth.
- Developed and validated **force field parameters** and simulation protocols for cellulose-hemicellulose systems, analyzing nanoscale mechanical responses under tensile and shear loading conditions.

Diffusion-based molecular design in flexible protein pockets (Collaborator)

- Contributed to methodology discussions for **YuelDesign**, a novel diffusion-based framework for structure-based drug design that models protein pockets as flexible, dynamic structures rather than rigid snapshots. (Wang et al., *Science Advances*, 2026)
- Contributed to methodology discussions on **YuelDesign**'s **SE(3)-equivariant diffusion framework** for drug design in flexible protein pockets, including the representation of pocket flexibility in three-dimensional molecular generation.
- Contributed to methodology discussions on how **YuelDesign** treats protein pockets as flexible, dynamic structures rather than rigid snapshots within its diffusion-based drug design framework, and how this modeling choice frames structure-based molecular generation.

GNN-based protein-ligand binding site prediction (Collaborator)

- Contributed to methodology discussions for **YuelPocket**, a **Graph Neural Network** for unified protein-small molecule binding site prediction, trained on the large-scale **PLINDER dataset**. (Wang & Dokholyan, *PNAS*, 2026)
- Contributed to methodology discussions on **YuelPocket**'s graph neural network formulation for protein-small molecule binding site prediction and its use of the large-scale **PLINDER dataset** for model training.
- Contributed to methodology discussions on the scope of **YuelPocket** as a graph neural network for protein-ligand binding site prediction, including how **PLINDER** supports the published model's training setting across protein-small molecule structures.

- **Dalian University of Technology** — Liaoning, China
  Graduate Research Assistant, Advisors: Prof. Shan Tang and Prof. Xu Guo — *Sep 2019 – Jun 2022*
  State Key Laboratory of Structural Analysis for Industrial Equipment

Data-driven constitutive modeling for brain tissue mechanics

- Developed **Artificial Neural Network (ANN)** constitutive models in **Python/MATLAB** to predict anisotropic mechanical behavior of brain tissue considering intravascular pressure, building upon the data-driven modeling framework of Tang, Z. et al. (*Defence Technology*, 2023).
- Integrated ANN models into **finite element (FE)** simulations in **Abaqus**, achieving **95% accuracy** while reducing computational cost by **80%** compared to traditional hyperelastic models.

Nanoscale water transport on defective graphene surfaces

- Investigated diffusion dynamics of water nanodroplets on double-vacancy graphene using **MD simulations** in **LAMMPS**. Analyzed constraining effects of defects on water mobility and wetting behavior. Published in *Applied Surface Science* (2021).

Teaching Experience

- **Penn State University** — State College, PA
  Teaching Assistant, Department of Engineering Science and Mechanics — *2022 – 2024*
  - TA for **EMCH 211 (Statics)** and **EMCH 212 (Dynamics)** over two academic years, leading recitation sessions, grading assignments, and holding office hours.

Publications

- **Li, J.**, Wang, J., & Dokholyan, N. V. (2026). ChironRNA: Steric Clashes Resolution in RNA Structures via E(3)-Equivariant Diffusion. *bioRxiv*, 2026-03.
- **Li, J.**, Wang, J., Ekambaram, S., & Dokholyan, N. V. (2026). ATLAS: Graph-based 3D RNA Motif Library Incorporating non-Watson-Crick Interactions. Under review, *Journal of Molecular Biology*; preprint on *bioRxiv*, 2026-02.
- L. Deng, **J. Li**, S. Tang, & Z. Guo (2021). Diffusion of water nanodroplets on graphene with double-vacancy: The constraining effects of defect. *Applied Surface Science*: 573.151235

Presentations

- **ATLAS: Graph-based 3D RNA Motif Library Incorporating non-Watson-Crick Interactions.** Presented as poster (Center for RNA Molecular Biology Symposium, Penn State Huck Institutes of the Life Sciences, May 2025) and oral presentation (RNA Club Seminar Series, Center for RNA Molecular Biology, Penn State University, 2025).