# Yonglan Liu (Permanent Resident)

Computational Drug Discovery | Physics + ML

Computational Chemistry • Cheminformatics • Computational Structural Biology • AI/ML • MD Simulation

Phone: 234-281-6739 | Email: liuyonglan04@gmail.com | LinkedIn: in/yonglan-liu/ | Github | Google Scholar

## Summary

Computational Chemist, Computational Structural Biologist, and AI Scientist with 10+ years of experience advancing drug discovery through physics-based molecular modeling, cheminformatics, structural bioinformatics, and machine learning. Proven track record of leading end-to-end modeling and informatics strategies—spanning docking, molecular dynamics, alchemical free energy perturbation (FEP), generative protein design, and AI-driven decision support—to guide lead identification, prioritization, and optimization across small-molecule, peptide, and protein-centered programs. Experienced scientific partner to medicinal chemists, structural biologists, and experimental teams, translating complex computational analyses into actionable design hypotheses that improve decision quality across iterative design–make–test–learn (DMTL) cycles.

## Experience

**Guidehouse** Bethesda, MD, USA

*Technical Senior Analyst* | *Computational Chemist* 03/2025 - Present

- Led modeling and informatics strategy for NIH-funded discovery programs, integrating structure-based docking, molecular dynamics simulations, alchemical free energy perturbation (FEP), and machine-learning–based developability prediction to guide compound prioritization and lead optimization decisions.
- Owned the architecture, validation, and reproducibility standards of OpenMM- and MBAR-based FEP pipelines, enabling quantitative protein–ligand binding affinity prediction and uncertainty-aware decision support in real medicinal chemistry cycles.
- Curated and analyzed target-focused chemical libraries using cheminformatics workflows to support iterative DMTL cycles, partnering with medicinal chemists to leverage free-energy–informed SAR analysis for hypothesis refinement and compound selection.
- Applied structure-based, ligand-based, and fragment-based design strategies to balance potency, selectivity, and developability constraints across multiple discovery programs.
- Developed a generative protein and peptide design workflow combining RFdiffusion, ProteinMPNN, and AlphaFold-Multimer to explore de novo binder concepts for target-focused discovery programs.
- Served as a scientific bridge between computation and experiment, translating FEP trends, MD-derived conformational insights, virtual screening results, and ML predictions into clear, actionable synthesis and experimental prioritization recommendations.

**National Cancer Institute, National Institutes of Health** Frederick, MD, USA

*Postdoctoral Fellow* | *Research Fellow* 05/2021 – 02/2025

- Independently led multiple structure- and mechanism-driven discovery efforts in cancer signaling, driving hypothesis formation, computational strategy, and interpretation across protein–ligand and protein–protein interaction systems.
- Developed reusable computational pipelines integrating MD-based conformational sampling with free-energy analyses (MM/GBSA, FEP) to enable systematic evaluation across multiple targets and ligand series.
- Identified cryptic and allosteric binding pockets via MD-driven conformational exploration, informing inhibitor and PROTAC design strategies.
- Applied AI-based protein structure prediction tools (AlphaFold, RoseTTAFold) to generate working models for proteins and protein–protein complexes in the absence of experimental structures, enabling downstream modeling and interaction analysis.

**The University of Akron** Akron, OH, USA

*Research Assistant* 01/2017 - 05/2021

- Built computational chemistry and biology pipelines to model protein–protein, protein–ligand, protein–membrane, and peptide interactions, supporting small-molecule and peptide-based therapeutic design.
- Applied molecular docking and MD simulations to characterize binding mechanisms and conformational behavior, guiding structure- and sequence-level optimization.
- Developed ML/AI models to predict sequence/structure–property relationships for peptides, small molecules, and biofunctional materials, working closely with experimental collaborators for validation.

## Selected Projects

**De Novo Protein and Peptide Design Pipeline for Targeted Protein Binding**

**RFdiffusion • ProteinMPNN • AlphaFold-Multimer • Generative AI**

- Designed and implemented an end-to-end computational pipeline for **de novo protein and peptide binder generation** targeting specific protein interfaces, integrating structure-conditioned diffusion models, sequence design, and structure prediction.
- Identified binding hotspots and anchor residues through structural interface decomposition, per-residue interaction energy analysis, and geometric pocket mapping to define diffusion conditioning sites.
- Applied RFdiffusion to generate backbone conformations conditioned on target binding sites and geometric constraints.
- Used ProteinMPNN for sequence optimization of generated backbones, producing stable, designable protein and peptide candidates.
- Performed complex structure prediction and binding assessment using AlphaFold-Multimer to evaluate interface quality, structural confidence, and interaction geometry.
- Implemented automated filtering based on pLDDT, interface metrics, and structural clustering to prioritize high-confidence designs.
- Built a reproducible, scriptable pipeline enabling iterative generative design loops for binder discovery and optimization.

**Machine Learning–Driven Peptide Design & Experimental Validation**

- Built ML models to identify and design self-assembling peptides with anti-amyloid properties, targeting amyloid-related diseases including Alzheimer's disease and type 2 diabetes.
- Represented peptide physicochemical properties using NNAAIndex-based descriptors and performed feature analysis to identify determinants of aggregation behavior.
- Trained and optimized SVM models (Python, GridSearchCV) to discriminate self-assembling versus non-assembling peptides.
- Collaborated with experimental teams to validate predictions, resulting in five confirmed peptides with strong anti-amyloid activity, including one patented peptide *(Authorized patent: ZL201410100022.5)*.

**Multimodal Active Learning Platform for Small Molecular Drug Design**

- Architected and delivered a physics-aware, ML-driven active learning platform integrating structure-based docking with multitask predictive modeling to guide hit discovery, hit-to-lead prioritization, and lead optimization across iterative design–make–test–learn (DMTL) cycles.
- Built production-grade chemical data curation and labeling pipelines, unifying RDKit, public databases, and internal assay data through rigorous molecular standardization, salt/tautomer normalization, deduplication, assay harmonization, and activity/ADME label construction with embedded quality-control analytics.
- Developed a scalable multimodal deep learning framework in PyTorch that fuses curated ADME features, molecular fingerprints, protein–ligand interaction descriptors, and pretrained 3D molecular graph embeddings, improving robustness and generalization on sparse, noisy discovery datasets.
- Established platform-level ML best practices, including strict data-splitting and leakage control, uncertainty-aware prediction, model monitoring, and iterative retraining, while systematically reintegrating experimental and structural feedback to improve prioritization quality and reduce low-value experimental cycles.

**End-to-End Alchemical Free Energy Perturbation (FEP) Pipeline for Ligand Binding Affinity Prediction**
**(**GitHub Link: openmm-rbfe | openfe-rbfe-dashboard)

- Designed and implemented an end-to-end alchemical FEP pipeline using OpenMM, openmmtools, and MBAR, enabling automated prediction of ligand binding affinities in real protein–ligand discovery projects.
- Built a reproducible workflow covering protein preparation, ligand mapping, alchemical transformations, solvation, equilibration, production MD, and free energy analysis, supporting consistent and auditable computational decision-making.
- Developed Python utilities and a lightweight dashboard to visualize ΔΔG values, uncertainty estimates, and simulation QC metrics, enabling rapid compound comparison and prioritization across ligand series in real design cycles.
- Applied the pipeline to real medicinal chemistry use cases, demonstrating its value in guiding structure-based lead optimization and reducing reliance on heuristic scoring functions.

## Education

**The University of Akron** Akron, OH
*PhD in Chemical Engineering* 01/2017 - 05/2021
**Chongqing University** Chongqing, China
*Bachelor of Science in Bioengineering* 09/2010 - 06/2014

## Skills

- **Programming & Scientific Computing**
  Python, R, Java, Fortran, HPC environments
- **Molecular Modeling & Simulation & Cheminformatics**
  RDKit, OpenMM, NAMD, CHARMM, Schrödinger, AutoDock Vina, PyMOL,
- **Machine Learning & AI**
  PyTorch, TensorFlow, DeepChem, HuggingFace, XGBoost, AlphaFold2, AlphaFold-Multimer, RoseTTAFold, DiffDock, ProteinMPNN, RFdiffusion

## Certification

- Data Scientist by *Dataquest*
- Natural Language Processing (NLP) with Attention Models, Natural Language Processing (NLP) with Probabilistic Models, Natural Language Processing (NLP) with Sequence Models, Natural Language Processing (NLP) with Classification and Vector Spaces, and Convolutional Neural Networks (CNN) by *DeepLearning.AI (Coursera)*

## APPENDIX

**Selected publications**

- Yonglan Liu, Wengang Zhang, Hyunbum Jang, Ruth Nussinov, mTOR variants activation discovers PI3K-like cryptic pocket, expanding allosteric, mutant-selective inhibitor designs, **J. Chem. Inf. Model**. 65, 2, 966–980 (2025)
- Yonglan Liu, Dong Zhang, Yijing Tang, Yanxian Zhang, Xiong Gong, Shaowei Xie, Jie Zheng, "Machine learning-enabled repurposing and design of antifouling polymer brushes" **Chem. Eng. J.**, 420: 129872 (2021)
- Yonglan Liu, Dong Zhang, Yijing Tang, Yanxian Zhang, Yung Chang, Jie Zheng, "Machine learning-enabled design and prediction of protein resistance on self-assembled monolayers and beyond", **ACS Appl. Mater. Interfaces**, 13: 11306-11309 (2021)
- Chen Chen, Yonglan Liu, Jin Zhang, Mingzhen Zhang, Jie Zheng, Yong Teng, Guizhao Liang, "A quantitative sequence-aggregation relationship predictor applied as identification of self-assembled hexapeptides", **Chemom. Intell. Lab. Syst.**, 145: 7-16 (2015)
- Yonglan Liu, Jin Zhang, Xiaohua Chen, Jie Zheng, Guixue Wang, Guizhao Liang, "Insights into the adsorption of simple benzene derivatives on carbon nanotubes", **RSC Adv.**, 4: 58036-58046 (2014)
- Dandan Huang, Yonglan Liu, Bozhi Shi, Yueting Li, Guixue Wang, Guizhao Liang, "Comprehensive 3D-QSAR and binding mode of BACE-1 inhibitors using R-group search and molecular docking", **J. Mol. Graph. Model.**, 45: 65-83 (2013)
- Yonglan Liu, Yueting Li, Bozhi Shi, Kan Zhong, Yiqiang Shao, Yafei Zeng, Dandan Huang, Guixue Wang, Guizhao Liang, "Using Topomer CoMFA and Surflex-dock to analyze 3D-QSAR of GSK-3β inhibitors and their mechanism of action", **SCIENTIA SINICA Chimica**, 43: 198-208 (2013)
- Yonglan Liu, Mingzhen Zhang, Hyunbum Jang, Ruth Nussinov, "The allosteric mechanism of mTOR activation can inform bitopic inhibitor optimization", **Chem. Sci.,** 15, 1003-1017 (2024).
- Yonglan Liu, Wengang Zhang, Hyunbum Jang, Ruth Nussinov, "SHP2 variants: the structural landscape and conformational dynamics can suggest the clinical phenotype, cancer or RASopathies", **Cell. Mol. Life Sci.**, 81:5 (2024)
- Mingzhen Zhang, Yonglan Liu, Hyunbum Jang, Ruth Nussinov, "Strategy toward Kinase-Selective Drug Discovery", **J. Chem. Theory Comput.**, 5: 1615-1628 (2023)
- Yonglan Liu, Mingzhen Zhang, Chung-Jung Tsai, Hyunbum Jang, Ruth Nussinov, "Allosteric regulation of autoinhibition and activation of c-Abl", **Comput. Struct. Biotechnol. J.**, 20: 4257-4270 (2022)
- Yonglan Liu, Hyunbum Jang, Mingzhen Zhang, Chung-Jung Tsai, Ryan Maloney, Ruth Nussinov, "The structural basis of BCR-ABL recruitment of GRB2 in chronic myelogenous leukemia", **Biophys. J**., 121: 2251-2265 (2022)
- Yanxian Zhang, Yonglan Liu (Co-first author), Yijing Tang, Dong Zhang, Huacheng He, Jiang Wu, Jie Zheng, "Antimicrobial alpha-defensins as multi-target inhibitors against amyloid formation and microbial infection", **Chem. Sci.**, 12: 9124 (2021)
- Baiping Ren, Yonglan Liu (Co-first author), Yanxian Zhang, Yongqing Cai, Xiong Gong, Yung Chang, Lijian Xu, Jie Zheng, "Genistein: A dual inhibitor of both Aβ and hIAPP peptides", **ACS Chem. Neurosci.**, 9: 1215-1224 (2018)
- Yonglan Liu, Jin Zhang, Xiaohua Chen, Jie Zheng, Guixue Wang, Guizhao Liang, "Insights into the adsorption of simple benzene derivatives on carbon nanotubes", **RSC Adv.**, 4: 58036-58046 (2014)