Talal Widatalla

talalw@stanford.edu • 323-630-8898 • San Francisco, CA

SUMMARY

Machine learning researcher at Arc Institute working in biological design and discovery, with interdisciplinary background in molecular biophysics as well as statistics and machine learning. Expertise in developing novel deep learning architectures for biological applications during PhD advised by Brian Hie, advancing diffusion-based protein generative models at Generate Biomedicines, as well as experience in building production ML systems in industry for molecular property prediction and hit-identification with Merck.

EDUCATION

Stanford University Expected 2027
PhD Student in Biophysics
NSF Graduate Research Felow
Stanford Graduate (SGF) Fellow

Johns Hopkins University May 2022
Bachelor of Arts in Biophysics
Bachelor of Science in Applied Math and Statistics

SELECTED WORK

- **Widatalla, T.**, Borah, A.A., King, S.H., Driscoll, C.L., Rafailov, R., Hie, B.L. "Aligning Protein Generative Models to Experimental Fitness with ProteinDPO." *Nature Methods* (2026). *In press.*
- Mille-Fragoso, L.*, Driscoll, C.*, Wang, J.*, Dai, H.*, **Widatalla, T.***, Zhang, J.*, Zhang, X., Rao, B., Feng, L., Hie, B., Gao, X. "Efficient generation of epitope-targeted de novo antibodies with Germinal" (2025). *Nature Biotechnology. In Press.*
- **Widatalla, T.***, Shuai, R.*, Huang, P-S, Hie, B. "Sidechain conditioning and modeling for protein sequence design with FAMPNN" (2025). *ICML.*
- **Widatalla, T.**, Rollins, Z., Cheng, A. "AbPROP: Language and Graph Deep Learning for Antibody Property Prediction" (2023). *ICML Workshop on Computational Biology.*
- Rollins, Z., **Widatalla, T.**, Cheng, A., Metwally, E. "AbMelt: Learning Antibody Thermostability from Molecular Dynamics" (2024). *Biophysical Journal.*
- Rollins, Z., **Widatalla, T.**, Waight, A., Cheng, A., Metwally, E. "AbLEF: Antibody Language Ensemble Fusion for Thermodynamically empowered property predictions" (2023). *Bioinformatics.*
- Waight, A., et al. (inc. **Widatalla, T.**) "A machine learning strategy for the identification of key in silico descriptors and prediction models for IgG monoclonal antibody developability properties" (2023). *mAbs.*

TECHNICAL SKILLS

- **Machine Learning:** GNNs, Language Models, Diffusion Models, Reinforcement Learning
- **Molecular Modeling:** Rosetta, Gaussian, MOE, MDAnalysis, GROMACS, Docking
- **Bioinformatics:** BLAST, HMMER, Clustal Omega, CD-HIT, SnapGene, UniProt
- **Programming:** Python, PyTorch, Linux, High Performance Computing, R, Matlab

EXPERIENCE

Graduate Student, Brian Hie Lab Jan 2024 – Present
Stanford University and the Arc Institute

- Implemented novel graph-neural network architecture and multi-modal training objective for full-atom protein sequence design method leading to SOTA performance across protein design benchmarks
- Developed approach for aligning structure-conditioned protein generative models with experimental fitness with offline reinforcement learning policy and led follow-up lab experiments for stability-enhanced viral immunogen design
- Contributed to first open-srouce generative framework *de novo* antibody binder design with 5-22% hit-rates
- Supervised undergraduate researchers on various projects in protein machine learning

Machine Learning Intern June 2025 – Sep 2025
Generate Biomedicines

- Expanded conditioning capabilities of a diffusion-based protein generative models for improved enzyme and binder design
- Conducted task formulation, architecture engineering, large scale multi-node training and evaluation
- Implemented custom loss functions and training tasks to instill new objectives in internal models

Data Scientist – AI/ML Aug 2022 – Sept 2023
Merck Research Laboratories

- Led development of production-ready ML system combining GNNs and language models for antibody property prediction, resulting in publication at ICML Workshop
- Designed and deployed end-to-end ML pipeline for antibody thermostability prediction using MD simulation features, work published in Biophysical Journal
- Architected high-throughput virtual screening system using active learning and structural deep learning, enabling rapid molecule screening at scale

Scientific Software Development Intern May 2022 – Aug 2022
OpenEye Scientific Software

- Developed software for mining structural properties from PDB protein structures
- Analyzed protein-interaction patterns to guide structure-based ligand design
- Evaluated protein modeling software through computational geometry analysis

Undergraduate Researcher Sept 2021 – May 2022
Johns Hopkins University, Dr. Jeffrey Gray Lab

- Evaluated zero-shot and few-shot fitness prediction capabilities of antibody language model trained on OAS database

SELECTED COURSEWORK

Generative Models • Computational Protein Design • Computational Chemistry • Probability & Statistics • Data Structures • Data Science • Applied Statistics • Optimization • Methods in Molecular Biophysics • Computational Biology