Logan Hallee

lhallee@udel.edu | logan@synthyra.com | linkedin.com/in/logan-hallee | Last updated November 2026

EDUCATION

PhD | University of Delaware, Bioinformatics Data Science | May 2026
Dissertation: *Representation Learning for Protein Annotation and Design*
Committee: Cecilia Arighi PhD, Mahdi Belcaid PhD, Li Liao PhD
Advisor: Jason Gleghorn PhD

BS | University of Delaware, Quantitative Biology | May 2022
Minored in Biochemistry

AAS | Delaware Technical Community College, Biological Sciences | May 2020

PROFESSIONAL EXPERIENCE

Synthyra | Cofounder and Chief Scientific Officer | July 2024 - Present

- Define protein-language-model strategy, commercialization, and team development for a Public Benefit LLC building molecular modeling tools for life-sciences research.
- Lead collaborative development of Atlas, a chemical language-model platform for proteome-scale protein-protein and protein-ligand interaction prediction, and Synthyra's iterative protein-design environment.
- Deliver web and API services that screen proteins and small molecules against whole proteomes, returning ranked interaction hypotheses, confidence scores, functional annotations, pathway enrichment, and structure models for experimental follow-up.
- Manage open-source protein language model projects whose models have been downloaded more than 3 million times on Hugging Face.
- Design sequence-scoring and mutation workflows for protein optimization; an EGFR-binder design reached a 630 pM dissociation constant against 3.33 nM for the cetuximab scFv control in third-party BLI validation, with six designs binding more tightly than that control.

RESEARCH EXPERIENCE

University of Delaware, Newark DE | Center for Bioinformatics & Computational Biology
Postdoctoral Research Fellow, Gleghorn Lab | May 2026 - Present

Accelerating Enzyme Engineering with Generative Protein Language Models

- Build high-quality datasets and predictive models with the CREATE BioFoundry for iterative protein design campaigns.

University of Delaware, Newark DE | Center for Bioinformatics & Computational Biology
Graduate Research Assistant, Gleghorn Lab | July 2022 - May 2026
Representation Learning for Protein Annotation and Design

- Invented Annotation Vocabulary: transformer-readable descriptions that improve protein representation learning.

- Produced first out-of-training-distribution *de novo* designed proteins with hallucinated domains matching prompts via BLAST hits and enrichment analysis.
- Architected protein diffusion models that generated novel proteins with picomolar predicted binding affinity.
- Designed Mixture of Experts extensions for converting pretrained language models into differentiated expert systems.
- Developed Protify, a modular platform for evaluating and fine-tuning protein language models across protein property prediction tasks.
- Created WINDOWZ, a computer vision pipeline for automated biomedical image segmentation from minimal training data and 3D reconstruction from 2D Z-stacks.

Indiana University, Virtual | Luddy School of Informatics, Computing, and Engineering
Machine Learning Consultant, Khomtchouk Lab | June 2023 - July 2024
Analysis of the Differential Influence of HDL in NHPI Populations

- Consulted on machine learning and statistical analysis of high-density lipoprotein-associated cardiometabolic outcomes in Native Hawaiian and Pacific Islander populations.
- Supported feature engineering, model evaluation, and biological interpretation of population-specific effects.

Wolfram Research, Virtual | Wolfram Winter School
Mentors: Stephen Wolfram PhD, James Boyd, Eric Parfitt | December 2022 - January 2023
Tetris for Proteins: Shape-Based Molecular Chemistry in the Wolfram Language

- Developed a shape-based computational abstraction for molecular chemistry and protein design in the Wolfram Language.
- Built interactive representations connecting molecular geometry, spatial packing, and protein structure-function reasoning.

University of Delaware, Newark DE | Mathematical Sciences Department
Research Project, Dr. Dominique Guillot | August 2021 - December 2022
Estimating the Graphical Structure of Climate Fields

- Studied graphical structure estimation for climate fields using applied mathematics and statistical dependency modeling.
- Worked at the interface of multivariate statistics, penalized regression, graph estimation, and geophysical data analysis.

University of Delaware, Newark DE | Biological Sciences Department
Undergraduate Researcher, Lam Lab | June 2021 - June 2022
Expanding the Domain-Based HSP90 Interactome for Various Applications for Cardiac Diseases

- Expanded a domain-based HSP90 interactome to support cardiac disease hypothesis generation and pathway analysis.
- Integrated protein domain knowledge, interaction evidence, and cardiac biology context for therapeutic target interpretation.

HONORS AND AWARDS

DARPA SBIR Direct-to-Phase II | 2027 - 2028
Principal Investigator of *Real-time multispecies host-pathogen interaction mapping for rapid target discovery*. Competitively selected for negotiation and possible award, September 2026. Synthyra main award with a University of Delaware subaward.

Innovation Delaware Fellows | 2026-2027

Brings together innovation and entrepreneurship ecosystem leaders to provide Delaware researchers critical knowledge, skillsets, and connections to successfully translate from the lab to the marketplace.

Institute for Engineering Driven Health Postdoctoral Researcher Award | 2026 - 2027

Funding support for outstanding postdoctoral researchers conducting translational work. This fellowship comes with training focused on entrepreneurial activities and bringing novel science and engineering to market. Funded by the NSF Accelerating Research Translation program number 2331440.

T32 Predoctoral Training Grant | 2025 - 2026
One-year predoctoral training grant from the University of Delaware Center for Bioinformatics and Computational Biology, supported by NIH NIGMS T32GM142603. The grant's mission is to support the strongest predoctoral candidates with multidisciplinary research training, cross-campus research infrastructure, team science skills, and workforce development.

Unidel Distinguished Graduate Scholars Fellowship | 2022 - 2027
Highly competitive five-year fellowship offered to doctoral students by the University of Delaware. Used to recruit exceptional doctoral students and provide them with intellectual and professional experiences that foster the development of disciplinary and community leaders.

William Clark Prize | 2022
In recognition of unusual talent and ability in mathematics among the graduating math seniors.

Math and Physics Department Outstanding Performance Award | 2020
Selected out of Delaware Technical Community College class of 2020.

PUBLICATIONS

Peer reviewed

Hallee, L; Peleg, T; Rafailidis, N; Gleghorn, J.P. Protein Language Models are Accidental Taxonomists. *BMC Bioinformatics*. 2026, 27, 188, doi.org/10.1186/s12859-026-06491-3, bioRxiv: doi.org/10.1101/2025.10.07.681002.

Hallee, L; Kapur, R; Patel, A; Gleghorn, J.P; Khomtchouk, B.B. Contrastive learning and mixture of experts enables precise vector embeddings in biological databases. *Nature Sci Rep*. 2025, 15, 14943, arXiv: doi.org/10.48550/arXiv.2401.15713.

Hallee, L; Khomtchouk, B.B. Machine learning classifiers predict key genomic and evolutionary traits across the kingdoms of life. *Nature Sci Rep.* 2023, 13, 2088, bioRxiv: doi.org/10.1101/2022.07.20.500846.

Roberts, R.J; **Hallee, L**; Lam, C.K. The Potential of Hsp90 in Targeting Pathological Pathways in Cardiac Diseases. *J. Pers. Med.* 2021, 11, 1373.

Sen, S; **Hallee, L**; Lam, C.K. The Potential of Gamma Secretase as a Therapeutic Target for Cardiac Diseases. *J. Pers. Med.* 2021, 11, 1294.

Publications submitted

Cotet, T-S; Krawczuk, I; Stocco, F; Ferruz, N; Gitter, A; Kurumida, Y; de Almeida Machado, L; Paesani, F; Calia, C. N; Challacombe, C. A; Haas, N; Qamar, A; Correia, B. E; Pacesa, M; Nickel, L; Subr, K; Castorina, L. V; Campbell, M. J; Ferragu, C; Kidger, P; **Hallee, L**; Wood, C. W; Stam, M. J; Kluonis, T; Ünal, S. M; Belot, E; Naka, A; Adaptyv Competition Organizers. (2025): Crowdsourced Protein Design: Lessons From the Adaptyv EGFR Binder Competition. In review at *Nature Methods*, bioRxiv: doi.org/10.1101/2025.04.17.648362.

Preprints

- Rafailidis, N; **Hallee, L**; Peleg, T; Horger, C; Haviland, L; Gleghorn, J.P. (2026): Protify: A Benchmarking Platform Revealing Task Dependence and Homology Leakage in Protein Language Model Evaluation. *bioRxiv*. Preprint. doi.org/10.64898/2026.11.04.764373
- Natoli, A; Reinhard, B; **Hallee, L**; Zhang, H; Piza-Taylor, K; Aung, F; Poznahovska, A; Gleghorn, J.P; Lucas, R; DiVita, M; Shen, Y. (2026): VPS34 regulation in pulmonary arterial vascular smooth muscle cells in pulmonary hypertension. *bioRxiv*. Preprint. doi.org/10.64898/2026.09.28.752142
- **Hallee, L**; Gleghorn, J.P. (2026): Dual Triangle Attention: Effective Bidirectional Attention Without Positional Embeddings. *arXiv*. Preprint. doi.org/10.48550/arXiv.2604.18603
- **Hallee, L**; Rafailidis, N; Bichara, D.B; Gleghorn, J.P. (2025): Diffusion Sequence Models for Enhanced Protein Representation and Generation. *arXiv*. Preprint. doi.org/10.48550/arXiv.2506.08293
- **Hallee, L**; Rafailidis, N; Horger, C; Hong, D; Gleghorn, J.P. (2024): Annotation Vocabulary (Might Be) All You Need. *bioRxiv*. Preprint. doi.org/10.1101/2024.07.30.605924
- **Hallee, L**; Rafailidis, R; Gleghorn, J.P. (2023): cdsBERT - Extending Protein Language Models with Codon Awareness. *bioRxiv*. Preprint. doi.org/10.1101/2023.09.15.558027
- **Hallee, L**; Gleghorn, J.P. (2023): Protein-Protein Interaction Prediction is Achievable with Large Language Models. *bioRxiv*. Preprint. doi.org/10.1101/2023.06.07.544109

Book Chapters

- **Hallee, L**; Jakubowski, H. Predicting Structure and Function of Biomolecules Through Natural Language Processing Tools. In *Fundamentals of Biochemistry Vol. I - Structure and Catalysis*, Chapter 4, The Three-Dimensional Structure of Proteins, section 4.13; Jakubowski, H.; Flatt, P., Eds.; Biology LibreTexts, 2026.

INVITED TALKS AND SEMINARS

Leveraging Annotation Vocabulary and Masked Diffusion for Sequence-based Protein Design

Department of Electrical and Computer Engineering Department Seminar Series, University of Delaware, 2025

Hallee, L; Rafailidis, N; Peleg, T; Bichara, D; Horger, C; Hong, D; Gleghorn, J.P.

Leveraging Annotation Vocabulary and Masked Diffusion for Sequence-based Protein Design

Invited talk, Cradle Bio, Zürich Switzerland, 2025

Hallee, L; Rafailidis, N; Peleg, T; Bichara, D; Horger, C; Hong, D; Gleghorn, J.P.

Leveraging Annotation Vocabulary and Masked Diffusion for Sequence-based Protein Design

Invited talk, Adaptyv Bio, Lausanne Switzerland, 2025

Hallee, L; Rafailidis, N; Peleg, T; Bichara, D; Horger, C; Hong, D; Gleghorn, J.P.

Speedrunning Protein Language Model Pretraining
Invited talk, Novonesis and Danish Technical University, Virtual, 2025
Hallee, L; Gleghorn, J.P.

High-level Protein Representation Learning with the Annotation Vocabulary
Seminar, Center for Bioinformatics and Computational Biology, University of Delaware, 2025
Hallee, L; Rafailidis, N; Horger, C; Hong, D; Gleghorn, J.P.

CONFERENCE PROCEEDINGS AND PRESENTATIONS

Protein Phosphatase 1 Catalytic Subunit Alpha Upregulation is Associated With AKT-Mediated Activation of Vacuolar Sorting Protein 34 in Pulmonary Arterial Smooth Muscle in Pulmonary Hypertension
Abstract submitted, New York State American Thoracic Society Seminar, 2026
Piza-Taylor, K; Aung, F; Natoli, A; Hallee L; Gleghorn J.P., Shen, Y.

Pulmonary arterial smooth muscle VPS34 activation is a shared molecular mechanism in adult pulmonary hypertension and congenital diaphragmatic hernia-associated pulmonary vascular remodeling
Abstract, American Heart Association's Scientific Sessions and Resuscitation Science Symposium, 2025
Published in *Circulation*, Volume 152
Gonyea, C; Natoli, A; Hallee, L; Gleghorn J.P; Shen, Y.

Akt-dependent Vacuolar Protein Sorting 34 Activation in Pulmonary Arterial Vascular Is a Shared Molecular Mechanism in Adult Pulmonary Hypertension and Congenital Diaphragmatic Hernia-associated Pulmonary Vascular Remodeling
Abstract, American Thoracic Society (ATS) International Conference, 2025
Published in *Am J Respir Crit Care Med*, Volume 211
Shen, Y; Gonyea, C.R; Hallee, L; Gleghorn, J.P.

Protify: A Low-Code Protein Property Prediction Tool
Poster, University of Delaware Bioinformatics Data Science Symposium, 2025
Rafailidis, N; Peleg, T; Hallee, L; Gleghorn, J.P.
Second place in poster competition.

Protify: A Low-Code Protein Property Prediction Tool - Subcellular Localization and Taxonomy Case Studies
Poster, Biomedical Engineering Society Annual Meeting, 2025
Peleg, T; Rafailidis, N; Hallee, L; Gleghorn, J.P.

Biomimetic Protein Sequence Generation and Representation with Annotation Vocabulary
Poster, University of Delaware Synthetic Biology Symposium, 2025
Hallee, L; Rafailidis, N; Horger, C; Hong, D; Gleghorn, J.P.
Second place in poster competition.

Targeting ER/SR HAX-1/Hsp90 Interaction to Enhance Cardiomyocyte Contractility and Survival upon Stress
Abstract, American Heart Association's Basic Cardiovascular Sciences Scientific Sessions 2025: Advances in Cardiovascular Science: From Discovery to Translation
Sen, S; Roberts, R; Wisner, H; Hallee, L; Lam, C.K.

Artificial intelligence-powered drug discovery: A case study on LD50 prediction and

molecular optimization
Poster, International Conference for High Performance Computing, Networking, Storage, and Analysis, 2024
Hallee, L; Roa, N; Rafailidis, N; Le, T; Horger, C; Singh, H; Ohashi, N; Lu, P.

A Supervised Machine Learning Approach Using Lung Vasculature to Classify Congenital Diaphragmatic Hernia Severity
Oral presentation, Biomedical Engineering Society Annual Meeting, 2024
Miller, M; Hallee, L; Gonyea, C; Gleghorn, J.P.

A Dynamic Time Warping-Based Approach for 2D Affine Image Alignment
Poster, Biomedical Engineering Society Annual Meeting, 2024
Wang, L; Hallee, L; Gleghorn, J.P.

HDL Cholesterol Confers Exceptionally Strong Cardioprotective and Metabolic Effects in Native Hawaiian and Pacific Islander Populations
Abstract, American Heart Association's Epidemiology and Prevention/Lifestyle and Cardiometabolic Health 2024 Scientific Sessions, 2024
Published in Circulation, Volume 149
Szatrowski, A; Sun, P; Lee, Y; Lord, G; Hallee, L; Hedges, J; Khomtchouk, B.B.

Windowed Biomedical Image Segmentation with Attentive-Recurrent-UNET
Poster, University of Delaware AI Center of Excellence Workshop, 2023
Hallee, L; Donzanti, M; Gleghorn, J.P.

WINDOWZ: 2D Segmentation and 3D Reconstruction of Biomedical Z-stacks
Oral presentation, University of Delaware Computer and Information Sciences Research Day, 2023
Hallee, L; Donzanti, M; Gleghorn, J.P.

SERVICE

Peer review

- *BMC BioData Mining*
- *BMC Genomics*
- *IEEE Transactions on Computational Biology and Bioinformatics*
- *Nature Communications Biology*
- *Nature Machine Intelligence*
- *Nature Scientific Reports*
- *The Protein Journal*

MENTORING EXPERIENCE

Tamar Peleg | University of Delaware, Biomedical Engineering major | February 2025 - Present
Democratizing molecular property prediction tools with open-source software
University of Delaware Summer Scholar

Tom Le | University of Delaware, Biomedical Engineering major | May 2023 - July 2025
Predicting taxonomy from protein primary sequences
University of Delaware INBRE Fellow

Colin Horger | University of Delaware, Biomedical Engineering major | May 2023 - May 2025

Exploring pooling methods on transformer latents
University of Delaware Summer Fellow, INBRE Fellow

Lindsey Wang | University of Delaware, Electrical Engineering and Applied Mathematics double major | August 2023 - February 2025
2D alignment and interpolation techniques for segmented Z-stacks
University of Delaware Summer Scholar, Winter Fellow, INBRE Fellow

Aaron Oster | University of Delaware, Computer Science major | January 2023 - February 2024
Enzyme commission number prediction with protein language models
AI Center of Excellence Spring Scholar, Summer Scholar, and Winter Fellow

John Bean | University of Delaware, Computer Science major | May 2023 - January 2024
WINDOWZ: An accessible machine learning pipeline for biomedical image segmentation and 3D reconstruction
INBRE Fellow

Nikolaos Rafailidis | University of Delaware, Biological Sciences major | May 2023 - December 2023
cdsBERT: Improving protein language model features with codon vocabulary extensions
McNair Summer Scholar

Jack Felipe | University of Delaware, Biomedical Engineering PhD student | August 2023 - October 2023
Automatic meshing from WINDOWZ point clouds
Rotating graduate student

William Sharp | University of Delaware, Computer Science major | May 2023 - August 2023
Text-to-text CLIP for enhanced protein semantic representations
University of Delaware Summer Scholar

PROFESSIONAL AFFILIATIONS

Sigma Xi Honor Society | 2021 - Present
Associate Membership

Phi Theta Kappa Honor Society | 2019 - Present
Full Membership