# Ryann M. Perez

302-423-2632 • ryann.m.perez@gmail.com • Philadelphia, PA • https://www.linkedin.com/in/ryann-perez/

## SUMMARY

Accomplished computational biochemist with over 5 years of experience translating *in silico* experimental techniques into real-world results. Accelerated adoption of generative AI and large language models (LLMs) in biological chemistry. Harnessed machine learning (ML), simulations, and statistical analysis in collaboration with cross-disciplinary teams to unravel complex molecular mechanisms. Seeking opportunities in building novel and impactful AI systems.

## EDUCATION

**PhD Candidate, Computational Biochemistry** *Expected March 2026*
University of Pennsylvania | GPA: 3.99
**Bachelor of Science, Chemistry (Minor: Biochemistry)** *May 2020*
University of Delaware | GPA: 3.72

## RESEARCH EXPERIENCE HIGHLIGHTS

**University of Pennsylvania** *September 2020 - Present*
**Graduate Researcher in the lab of Dr. E. James Petersson**

- Crafted custom LLMs to achieve state-of-the-art performance on biologic aggregation and protein stability; important findings that streamline the property prediction of proteins. Scaled LLM for inference on 64 million datapoints
- Trained students to use a novel LLM system for biological chemistry question and answer tasks which demonstrably improved learning experiences. This technology has been adopted by faculty for recurrent use in all future classes
- Constructed machine learning pipelines to predict ligand affinity toward alpha-synuclein fibrils
- Enhanced productivity by creating custom Python software packages to streamline experimental analysis, drug discovery, and protein simulations

## SELECTED AWARDS & PRESENTATIONS

- **Predoctoral Individual National Research Service Award (F31)**: Awarded from the National Institute of Health (NIH) for an original proposal describing and prototyping an amyloid polymorphism determination assay *September 2024*
- **University of Pennsylvania Dean's Scholar**: Given to 20 students in the School of Arts and Sciences for outstanding academic achievement and intellectual promise *April 2024*
- **Lectures on Generative AI in Chemistry:** Designed and led seminars attended by over 50 professionals and scientists on utilizing ChatGPT in chemistry workflows, leading to increased awareness and implementation of generative AI within the Chemistry Department. *May 2025*
- **Invited Lecture on AI in Biochemistry, Temple University**: Generated and presented an original deep learning lecture for an audience who were novices in biochemistry *April 2024*
- **Poster Presentations**: ChemBio in the Hub – *September 2025* – American Society of Biochemistry and Molecular Biology – *March 2025* – American Chemical Society – *August 2023 & August 2025*

## TEACHING EXPERIENCE

**University of Pennsylvania**
**Teaching Assistant in Biological Chemistry** *Spring Semesters 2023 – 2024*

- Mentored over 60 undergraduate and graduate students in biological chemistry. Crafted one-hour lectures on AI in biochemistry and presented to the class

## TECHNICAL SKILLS

- **Programming & ML:** Classical ML, deep learning (MLP, CNN, GCN, encorders/decoders, autoencoders, transformers, large language models or LLMs, diffusion models, contrastive learning), fine tuning, parameter efficient fine-tuning (PERT), domain adaptation, retrieval augmented generation (RAG), multi-GPU and

distributed training, exploratory data analysis (EDA), dimensionality reduction, unsupervised learning, dataset curation, class imbalances, experimental biases, statistical analysis. Proficient in Python (PyTorch/Tensorflow, HuggingFace, scikit-learn, NumPy, Pandas). Amazon Web Services. GitHub & version control.

- **Computational Chemistry:** Custom virtual ligand screening. Simulations (Rosetta/PyRosetta). PyMOL. ADMET prediction. Structure activity relationships (SAR). Machine learning in biological contexts (Stability, solubility, aggregation).
- **Chemistry:** Organic synthesis, LC-MS, HPLC, NMR, peptide synthesis, flash chromatography.
- **Biology:** Protein expression/purification, SDS-PAGE, FPLC, MALDI-TOF MS, enzymatic assays, high throughput assay design and execution, fluorescence polarization (FP).

## SELECTED PUBLICATIONS

- **Perez, R. M.**; Shimogawa M.; *et al.* Large Language Models for Education: ChemTAsk -- An Open-Source Paradigm for Automated Q&A in the Graduate Classroom DOI: https://arxiv.org/abs/2502.00016 *Comput. Educ.: Artif. Intel.* ***Accepted***
- Li, X.; **Perez, R. M**.; Mach, R. H.; Giannakoulias, S.; Petersson, E. J. Machine Learning Prediction of Multiple Distinct High-Affinity Chemotypes for α-Synuclein Fibrils *Chem. Commun.* **2026**. DOI: 10.1039/D5CC06228D
- Li, X.; **Perez, R. M.**.; *et al.* Accurate Prediction of Protein Tertiary and Quaternary Stability Using Fine-Tuned Protein Language Models and Free Energy Perturbation. *Int. J. Mol. Sci.* **2025**, *26*, 7125. DOI: 10.3390/ijms26157125
- **Perez, R. M**.; Li, X.; Petersson, E. J.; Giannakoulias, S. AggBERT: Best in Class Prediction of Hexapeptide Amyloidogenesis with a Semi-Supervised ProtBERT Model. *J. Chem. Inf. Model.* **2023** DOI: 10.1021/acs.jcim.3c00817
- All publications may be found at: scholar.google.com/citations?user=nOgqdusAAAAJ&hl=en