

RICHARD SHUAI

richardshuai98@gmail.com | 818-398-1883 | www.linkedin.com/in/richard-shuai | <https://richardshuai.github.io/>

EDUCATION

Stanford University

Ph.D. Biophysics

Expected: June 2027

GPA: 3.94 / 4.0

University of California, Berkeley

B.A. Computer Science & B.A. Statistics

May 2022

GPA: 3.99 / 4.0

RESEARCH EXPERIENCE

Stanford University (Advisor: Po-Ssu Huang)

Jan. 2021 – June 2023

PhD Candidate

Exploring methods for all-atom protein design. Developed FAMPNN, a full-atom protein sequence design method relying on per-token Euclidean diffusion for sidechain generation. Developing diffusion- and flow-based methods for improved protein backbone generation and motif scaffolding.

Johns Hopkins University (Advisor: Jeffrey Gray)

Jun. 2021 – Sept. 2022

NSF REU Research Fellow

Used generative language modeling to design synthetic antibody libraries with favorable developability properties. Trained a ~12.8M parameter GPT-based transformer on ~558M antibody sequences. Sampled novel antibody sequences, measured properties for therapeutic developability, and verified design foldability in silico via AlphaFold2 structural predictions.

UC Berkeley / UCSF (Advisors: Nilah Ioannidis, Gabriel Loeb)

Jan. 2021 – June 2023

Research Assistant

Used and interpreted CNN-based deep learning models with scATAC-seq data to prioritize variants affecting chromatin accessibility in kidney tissue. Trained models and evaluated performance on predicting allelic imbalance. Analyzed first-layer convolutional filters and in silico saturation mutagenesis outputs to discover transcription factor binding motifs.

UC Berkeley (Advisor: Laura Waller)

Nov. 2019 – Mar. 2021

Research Assistant

Used deep learning for real-time image reconstruction from convolutions with a spatially-variant point spread function. Proposed and trained a U-Net with learnable Wiener deconvolutions, achieving ~1000x speedup against existing iterative algorithms with similar reconstruction quality.

UC Berkeley / UCSF (Advisor: Reza Abbasi-Asl)

Jan. 2021 – Current

UC Berkeley URAP Apprentice

Developed statistical pipelines to associate multiple sclerosis patient genotypes and RNA-seq data with their phenotypes. Performed exploratory analysis of gene expression data using DESeq2 and clustering algorithms.

EMPLOYMENT

AbSci, AI Scientist Intern (Advisor: Joshua Meier)

April 2022 – Sept. 2022

Developed and trained models based on language models and graph neural networks to design and optimize antibodies for binding to a specific target. Wrote data processing pipelines for training models while ensuring rigorous evaluation of model generalization. Coordinated with the wet lab team to experimentally validate designed antibodies.

Amazon, Software Engineering Intern

May 2020 – Aug. 2020

Developed pricing configuration page for Amazon Pricing Team and external business users for competitive price analysis with Agile development. Expanded backend functionality within Java Spring Framework to add key

features. Revamped previous UI with React Redux, supporting 8 month phase-out of previous interface.

Abiocode, Inc., Lab Technician

May 2018 – Aug. 2018

Prepared Arabidopsis sample protein extracts for gel electrophoresis. Performed SDS-PAGE and Western Blotting experiments, using basic lab techniques such as pipetting and centrifugation.

WORKSHOPS & PRESENTATIONS

- **Richard W. Shuai**, Jeffrey A. Ruffolo, Jeffrey J. Gray. “Generative Language Modeling for Antibody Design.” *NeurIPS Workshop on Machine Learning for Structural Biology (MLSB)*, poster (2021).
- **Richard W. Shuai**, Jeffrey A. Ruffolo, Jeffrey J. Gray. “Generative Language Modeling for Antibody Design.” *RosettaCON*, poster (2021).
- **Richard Shuai**, Connie Huang, Parth Baokar, Ryan Chung, Ruchir Rastogi, Pooja Kathail, Nilah Ioannidis. “Personal transcriptome variation is poorly explained by current genomic deep learning models.” *American Society of Human Genetics*, (2023).
- Pooja Kathail, **Richard Shuai**, Ryan Chung, Chun Jimmie Ye, Gabriel Loeb, Nilah M. Ioannidis. “Assessing the utility of genomic deep learning models for disease-relevant variant effect prediction.” *ICML Workshop on Computational Biology*, spotlight (2022).
- Ayesha Bajwa, Ruchir Rastogi, Pooja Kathail, **Richard W. Shuai**, Nilah Ioannidis. "Characterizing uncertainty in predictions of genomic sequence-to-activity models." *Machine Learning in Computational Biology* (2024).

PREPRINTS & PUBLICATIONS

- **Richard W. Shuai***, Talal Widadalla*, Brian L. Hie[†], Po-Ssu Huang[†]. "Sidechain conditioning and modeling for full-atom protein sequence design with FAMPNN." *International Conference on Machine Learning* (2025).
- Alexander E Chu, Jinho Kim, Lucy Cheng, Gina El Nesr, Minkai Xu, **Richard W. Shuai**, Po-Ssu Huang. "An all-atom protein generative model." *Proceedings of the National Academy of Sciences* (2024).
- **Richard W. Shuai***, Jeffrey A. Ruffolo*, Jeffrey J. Gray. “Generative Language Modeling for Antibody Design.” *Cell Systems* (2023).
- Jeffrey Ruffolo, **Richard Shuai**, Jeffrey J Gray. "Generative Language Models and Related Aspects for Peptide and Protein Sequence Design." *US Patent App. 18/715,967* (2025).
- Amir Shanehsazzadeh, Sharrol Bachas, Matt McPartlon, George Kasun, John M Sutton, Andrea K Steiger, **Richard Shuai**, ..., Joshua Meier. "Unlocking de novo antibody design with generative artificial intelligence." *bioRxiv* (2023).
- Connie Huang*, **Richard Shuai***, Parth Baokar*, Ryan Chung, Ruchir Rastogi, Pooja Kathail, and Nilah Ioannidis. "Personal transcriptome variation is poorly explained by current genomic deep learning models." *Nature Genetics* (2023).
- Gabriel Loeb, Pooja Kathail, **Richard Shuai**, Ryan Chung, Rein Grona, Sailaja Peddada, Volkan Sevim, Scot Federman, Karl Mader, Audrey Chu, Jonathan Davitte, Shawn Shafer, Laralynne Przybyla, Radu Rapiteanu, Nilah Ioannidis, Jeremy Reiter. “Variants in tubule epithelial regulatory elements mediate most heritable differences in human kidney function.” *Nature Genetics* (2024).
- Pooja Kathail, **Richard W. Shuai**, Ryan Chung, Chun Jimmie Ye, Gabriel B. Loeb, Nilah Ioannidis. “Current genomic deep learning models display decreased performance in cell type-specific accessible regions.” *Genome Biology* (2023).
- Kyrollos Yanny*, Kristina Monakhova*, **Richard W. Shuai**, Laura Waller. “Deep learning for fast spatially-varying deconvolution.” *Optica* 9, 96-99 (2022).
- **Richard W. Shuai***, Kyrollos Yanny*, Kristina Monakhova, Laura Waller. “MultiWienerNet: Deep Learning for Fast Shift-Variant Deconvolution.” *Computational Optical Sensing and Imaging*, pages CTh5A–5. Optical Society of America (2021).

HONORS & AWARDS

- NSF Graduate Research Fellowship (2024 - Present)

SKILLS

Programming languages: Python, Java, R, C, Javascript

Libraries: PyTorch, Tensorflow, Keras, HuggingFace, NumPy, pandas, SciPy

Molecular modeling: PyRosetta, PyMOL

Relevant coursework: Machine Learning, Deep Learning, Computational Methods in Genomics and Medicine, Artificial Intelligence, Optimization Models, Probability and Random Processes, Data Structures, Linear Algebra, Linear Modeling