

Adit Shah

925-997-9055 | aditshah00@gmail.com | [linkedin.com/in/aditshah00](https://www.linkedin.com/in/aditshah00) | github.com/ashah03

EDUCATION

University of California, Berkeley

Berkeley, CA

B.S. Electrical Engineering & Computer Sciences (GPA: 3.98)

Aug. 2021 – May 2025

Selected Coursework: Machine Learning, AI Meets Bio & Chem, Algorithms for Computational Biology, Computational Genomics, Convex Optimization, Probability Theory & Random Processes, Statistical Mechanics, Classical Mechanics, Electricity & Magnetism, Discrete Mathematics, Computer Architecture, Artificial Intelligence, Operating Systems

EXPERIENCE

Columbia University

New York, NY

Research Associate – AlQuraishi Lab

Aug. 2026 – Present

- Developing a hypernetwork-based method to steer AlphaFold3 with experimental data such as cryo-EM, SAXS, and chemical probing to extend accurate structure determination to flexible proteins, RNA, and complexes
- Training purpose-built encoders that leverage the structure of each experimental data type, such as the Fourier-space representation of cryo-EM maps, to feed a shared, modality-agnostic conditioning framework
- Recovering up to 77% of idle GPU capacity on Columbia's shared cluster by building a simulator of past usage to test scheduling policies and collaborating with PIs and Columbia IT on implementation
- Building a foundation in biophysics by reading *Physical Biology of the Cell* and participating in the lab's focus group on conformations, including presenting literature on generative models for sampling transition pathways

Until Labs

South San Francisco, CA

Research Engineer

Jun. 2024 – Jul. 2026

- Trained models to predict small-molecule cryoprotectant efficacy and toxicity from electron densities, foundation model embeddings, and molecular descriptors, raising the hit rate in lab screens for ice inhibition
- Created a distributed DFT pipeline computing electron densities and thermodynamic properties for 220K small molecules in water, using spot instances with fault-tolerant job queuing to cut compute costs by 90%
- Built company's scientific data platform from scratch in collaboration with leadership and scientists across teams, providing templated analysis pipelines and tracking 1,200+ experiments and biological entities

Broad Institute of MIT and Harvard

Cambridge, MA

Research Intern – Imaging Platform

Jan. 2024 – May 2024

- Found a flaw in drug discovery benchmarks where many assays were trivially predicted by cell count rather than finer morphological features, leading to a co-authored publication in *Nature Communications*
- Built modular evaluation framework comparing cellular imaging, chemical structure, and gene expression for predicting biochemical assays, supporting flexible combinations of embeddings and models
- Established SkyPilot cloud infrastructure and GPU resource pipelines adopted by the broader lab

Berkeley Artificial Intelligence Research (BAIR)

Berkeley, CA

Undergraduate Researcher – Berkeley Speech Group

Aug. 2023 – Feb. 2024

- Curated and preprocessed datasets across three silent-speech recording modalities to demonstrate generalization of streaming neural speech decoding architectures
- Co-authored a *Nature Neuroscience* paper from an active clinical trial restoring speech to paralyzed individuals

Undergraduate Researcher – CIRCLES Consortium

Dec. 2021 – May 2023

- Trained and evaluated RL controllers for autonomous vehicles to reduce highway traffic congestion, including observation space design, reward shaping, and training curricula
- Demonstrated >15% improvement in fuel efficiency in simulation with just 4% autonomous vehicle penetration, leading to a co-first-authored publication at IEEE ITSC
- Deployed algorithm on 100+ vehicles for a week-long field test in Nashville, the largest autonomous vehicle traffic experiment to date, resulting in two *IEEE Control Systems* papers

EPFL (École Polytechnique Fédérale de Lausanne)

Lausanne, Switzerland

Summer@EPFL Research Fellow – MLBio Lab

May 2023 – Aug. 2023

- Developed a modular framework benchmarking meta-learning approaches to few-shot learning on biomedical tasks including drug response prediction, antibiotic resistance, and protein function annotation

SELECTED PROJECTS & INITIATIVES

- Machine Learning for Biology** | *Course Projects & Presentations* Aug. 2024 – May 2025
- *Fusion Protein Prioritization & Design*: Classified oncogenic fusions from protein language model embeddings at 0.877 AUC, and experimented with steering a LoRA-tuned protein diffusion model via a cross-attention adapter
 - *Chromatin State Prediction*: Extended the DeepSEA genomic CNN with dilated convolutions, attention, and longer input sequences, finding that a 2,000 bp context window improved performance
 - *Protein Foundation Models*: Presented a three-hour literature review of ESM3 to a graduate course, covering how tokenizing structure with a VQ-VAE enables joint generative modeling of protein sequence, structure, and function
- Neurotech @ Berkeley** | *Software Division Lead* Aug. 2022 – Jan. 2024
- Led a division of 30 Berkeley students working on projects related to software, ML, and neuroscience
 - Elected President for organization of 120 members
- CS 61B: Data Structures** | *Undergraduate Student Instructor & Tutor* Jun. 2022 – Jun. 2023
- Taught discussion and laboratory sections with 50+ students to teach concepts such as object-oriented programming, asymptotic analysis, hashing, sorting, and graph algorithms
 - Hosted office hours, graded student assignments/exams, and developed course curriculum (assignments and projects) for 1,500+ students

PUBLICATIONS

- Counting cells can accurately predict small-molecule bioactivity benchmarks.** S. Seal, W. Dee, **A. Shah**, . . . , S. Singh, O. Spjuth, A. Bender, A.E. Carpenter. *Nature Communications*, 2026.
- A streaming brain-to-voice neuroprosthesis to restore naturalistic communication.** K.T. Littlejohn, C.J. Cho, J.R. Liu, A.B. Silva, B. Yu, V.R. Anderson, C.M. Kurtz-Miott, S. Brosler, A.P. Kashyap, I.P. Hallinan, **A. Shah**, A. Tu-Chan, K. Ganguly, D.A. Moses, E.F. Chang, G.K. Anumanchipalli. *Nature Neuroscience*, 28, pp. 902–912, 2025.
- Traffic control via connected and automated vehicles (CAVs): An open-road field experiment with 100 CAVs.** J.W. Lee, H. Wang, K. Jang, N. Lichtlé, A. Hayat, M. Bunting, A. Alanqary, W. Barbour, Z. Fu, X. Gong, G. Gunter, S. Hornstein, A.R. Kreidieh, M.W. Nice, W.A. Richardson, **A. Shah**, E. Vinitzky, F. Wu, . . . , A.M. Bayen. *IEEE Control Systems*, 45(1), pp. 28–60, 2025.
- Reinforcement learning-based oscillation dampening: Scaling up single-agent reinforcement learning algorithms to a 100-autonomous-vehicle highway field operational test.** K. Jang, N. Lichtlé, E. Vinitzky, **A. Shah**, M. Bunting, M. Nice, B. Piccoli, B. Seibold, D.B. Work, M.L. Delle Monache, J. Sprinkle, J.W. Lee, A.M. Bayen. *IEEE Control Systems*, 45(1), pp. 61–94, 2025.
- Traffic smoothing controllers for autonomous vehicles using deep reinforcement learning and real-world trajectory data.** N. Lichtlé*, K. Jang*, **A. Shah***, E. Vinitzky*, J.W. Lee, A.M. Bayen. *2023 IEEE International Conference on Intelligent Transportation Systems (ITSC)*, pp. 4346–4351, 2023.

*Equal contribution

AWARDS & HONORS

- High Honors in Electrical Engineering and Computer Sciences** | *UC Berkeley* 2025
- Regents' & Chancellor's Scholar** | *UC Berkeley* 2021 – 2025
- Summer@EPFL Research Fellowship** | *EPFL* 2023

TECHNICAL SKILLS

Languages: Python, Java, C/C++, Rust, Kotlin, JavaScript

ML Frameworks: PyTorch, Scikit-Learn, XGBoost, OpenCV, Stable Baselines, OpenAI Gym

ML Methods: CNNs, Reinforcement Learning (PPO), Few-Shot Learning, VQ-VAE

Computational Science: OpenFold3, Protein Language Models (ESM), RDKit, ORCA, PyMOL

Data: NumPy, SciPy, Pandas, Polars, DuckDB, PostgreSQL, SQLAlchemy, Marimo

Infrastructure & Tools: AWS (SQS, S3, EC2 Spot), Azure, Modal, SkyPilot, Ray, SLURM, Docker, Git, gRPC