# Seyone Chithrananda

+1 510 990 7557 | seyonec@berkeley.edu | Website | GitHub | Google Scholar | Berkeley, California

## Education

**University of California, Berkeley** — Berkeley, California
*Bachelors of Arts in Computer Science* — *August 2021 – May 2025*

**Computer Science:** Reinforcement Learning (graduate course), Efficient Algorithms and Intractable Problems, Intro to Machine Learning, Artificial Intelligence, Computational Methods in Genomics, Computer Security, Machine Structures, Data Structures, Structure and Interpretation of Computer Programs
**Math/Stats/Biology coursework**: Analysis, Therapeutics Discovery and Development (chemical biology of modern biologics/SMDs, assay dev., etc.), Computational Functional Genomics (single-cell graduate course), Convex Optimization, Discrete Math and Probability Theory, Designing Information Devices and Systems I & II (linear algebra, circuit analysis, signal processing), Single/Multivariable Calculus.

## Experience

**Innovative Genomics Institute, Laboratory of Prof. Jennifer Doudna** — Berkeley, California
*Undergraduate Researcher* — *January 2023 – Present*

- Emphasis on advancing research encompassing frameworks for dealing with uncertainty in protein design, model-guided sequence design, de-biasing of complex multi-modal biological data, and remote homology detection.
- Developed a novel technique bridging conformal prediction for statistically-robust protein functional annotation, with applications to metagenomics, enzyme classification, and structural alignment in collaboration with Dr. Anastasios Angelopoulos and Prof. Michael I. Jordan's lab (in print, Nature Communications).
- Developed new datasets and minimized structure-aware generative models for the design of thermostable ribosomal RNA, conferring increased thermostabiliy to the *E. coli* ribosome. (in print, Nature Communications).
- Co-first author on protein semantic similarity search method for RNA-Guided endonuclease discovery *(Computational Biology, ICML 2023).*

**Arc Institute, Laboratory of Evolutionary Design (Prof. Brian Hie)** — Palo Alto, California
*Visiting Undergraduate Researcher* — *June 2024 – Sept 2024*

- Summer research internship, advised by Prof. Brian Hie, working on techniques for controllable DNA sequence generation, with the goal of enabling more programmable, top-down genetic structure from genome tracks.
- Training DNA language models that design genetic circuits based on conditioned positional annotations, with robust evals for the evaluation of de novo designed synthetic DNA sequences, measuring steerability of models.

**Microsoft Research** — Cambridge, Massachusetts
*Research Intern* — *May 2023 – August 2023*

- Advised by Dr. Kevin Yang and Dr. Judith Amores, developing machine learning models for olfaction.
- Developed multi-modal architecture for predicting protein-ligand interactions involving olfactory receptors to map combinatorial coding, using geometric learning and protein language models (under review, *Cell Systems*).
- Developed technique that can predict a molecules odor perception, using information about the olfactory receptors it activates through a sparse coding-inspired modeling methodology. Provisional patent awarded for method.

**Dyno Therapeutics** — Cambridge, Massachusetts
*ML Design Intern* — *May 2022 – August 2022*

- Researching methods for viral protein AAV capsid design, using sequence-to-function graph and sequence-based models (transformers, GNNs, etc). a16z, Google Ventures backed Church Lab startup, with 120M Series A.
- Developed generative structure-to-sequence models to propose high-scoring variant sequences, and examined performance at standard protein redesign, handling epistatic interactions, indels, and at predicting binding on experimentally-validated and in-silico fitness landscapes *(ML in Structural Biology, NeurIPS 2022).*

**Nurix Therapeutics** — San Francisco, California (remote)
*Computational Chemistry intern* — *May 2021 – Aug 2021*

- Developed graph-based and evolutionary deep learning methods for DNA-encoded library (DEL) and scaffold-based design, to do multi-objective inhibitor & linker design while accounting for noisy experimental data.

- Deployed multiple classification and regression models for screening molecules within core DEL-ML platform. Implemented message-passing, graph convolutional neural networks for binding affinity, ADME-tox modeling.

**University of Toronto, Laboratory of Prof. Alan Aspuru-Guzik** Toronto, Canada
*Research Student* *April 2020 – April 2021*

- Co-developer of SELFIES v1.0, a 100% robust molecular string representation for machine learning models. Developed depth-first graph traversal algorithm and dearomatization code for v1 release. Downloaded 9K+ times.
- Published a review paper, highlighting statistical methods for uncertainty estimation in ML for property prediction. *(Expert Opinions in Drug Discovery, 2021)* .
- Developed large-scale foundation models for chemistry with open-source organization DeepChem (ChemBERTa). Models have 12M+ API calls, 450+ citations *(ML for Molecules, NeurIPS 2020)*.

## Publications

**S Chithrananda**, J Amores, KK Yang. Mapping the combinatorial coding between olfactory receptors and perception with deep learning. In review, *Cell Systems*, (2024). bioRxiv.

R Boger, **S Chithrananda**, AN Angelopoulos, PH Yoon, MI Jordan, JA Doudna. Functional protein mining with conformal guarantees. In print, *Nature Communications*, (2024). bioRxiv.

Y Shulgina*, M Trinidad*, C.J. Langeberg*, H Nisonoff*, **S Chithrananda***, ..., JA Doudna, J.H. Cate*. RNA language models predict mutations that improve RNA function. In print, *Nature Communications*, (2024). bioRxiv.

R Boger*, AX Lu*, **S Chithrananda***, ..., P Abbeel, JA Doudna. TOPH: Adapting A Contrastive Question-Answering Framework for Protein Search. *ICML Workshop on Computational Biology*, (2023). Paper

J Chan, **S Chithrananda**, D Brookes, S Sinai. A Benchmark Framework for Evaluating Structure-to-Sequence Models for Protein Design. *NeurIPS ML for Structural Biology Workshop* (2022).

W Ahmad, E Simon, **S Chithrananda**, G Grand, B Ramsundar. ChemBERTa-2: Towards Chemical Foundation Models. *ELLIS ML for Molecules Workshop* (2021), ArXiv

**S Chithrananda**, G Grand, B Ramsundar. ChemBERTa: Large-Scale Self-Supervised Pre-training for Molecular Property Prediction. *NeurIPS 2020 ML for Molecules workshop* (2020), ArXiv

A Nigam, R Pollice, M.F. Hurley, ... **S Chithrananda**, V Voelz, A Aspuru-Guzik. Assigning Confidence to Molecular Property Prediction. *Expert Opinions in Drug Discovery, Taylor and Francis* (2021), Journal Article

## Awards & Achievements

**New Sciences Computational Life Sciences Fellowship** | *Website*

- Awarded grant funding (summer 2024, $10,000) for generative sequence design research with Brian Hie at Arc Institute, incorporating computational analyses that are well-informed or grounded by experimental data.

**Accel Scholars, 2024-25 Cohort** | *Website*

- The Accel Scholars program empowers undergraduate EECS students at UC Berkeley through unparalleled access to the tech industry, personalized mentorship, and industry-relevant curriculum. The program is a joint venture between venture capital firm Accel and UC Berkeley's EECS department.

**Masason Foundation Scholar** | *Website*

- Support from Masayoshi Son's foundation (founder of Softbank Group). Provided scholarships for research and tuition at UC Berkeley.

**Emergent Ventures Fellow** | *Press*

- Emergent Ventures is a low-overhead fellowship and grant program that supports entrepreneurs and brilliant minds with highly scalable, "zero to one" ideas for meaningfully improving society.
- Awarded two grants, one in March 2020 for research in computational chemistry, and a second to support living expenses for undergraduate studies at UC Berkeley.

**Re-Work Young Researcher:** | *Talk*

- Delivered talk at conference in front of over 1000 attendees on independent research in computational biology.

**D.E. Shaw Undergraduate Research Fellowship** | *Website*

- Fellowship program for undergraduates in computational drug discovery.

## Scientific Talks

AI in Biomedicine seminar: Annotating the proteome with conformal guarantees. Stanford University. Jan 2025

Towards mapping the combinatorial coding of olfaction. Osmo Labs (Google Brain spinout). Oct 2024

Rising Stars Seminar: Understanding the Combinatorial Code of Smell. Alaa Lab, UC Berkeley. Sept 1, 2023 | *Talk*

Scientific Machine Learning Webinar Series: ChemBERTa: Large-Scale Self-Supervised Pretraining for Molecular Property Prediction. Carnegie Mellon University. Feb 18, 2021 | *Talk*

Bay Area Machine Learning Symposium : ChemBERTa: Large-Scale Self-Supervised Pretraining for Molecular Property Prediction. Remote. Oct 15, 2020 | *Talk*

Re-Work Deep Learning Summit: Deep Learning to Understand Gene Expression. Montreal, Canada. Nov 3, 2019 | *Talk*

## Service

**Reviewing:** NeurIPS Workshop on ML for Structural Biology 2023, 2024

**Co-organizer: Berkeley BioML Seminar Series** | Bringing leaders at the cutting-edge of computation and biology in industry, academia and policy to Berkeley. Over 600 attendees during 23-24 seminar series, and sponsorship from Pillar VC. | *Overview* | 2023-24

**Machine Learning at Berkeley; VP of Research** Supporting Berkeley undergrads to do meaningful research, through organizing research talks, reading groups, project mentorship, & organizing seminars. Managed $100,000 in club funds, committee of 12 undergrads to create a platform for students and Berkeley faculty to collaborate on computational research, offer funding to attend conferences, and publish their work. | *Website* | 2021-2024

## Skills

**Programming:** C, C++, Java, Python

**Frameworks:** Pytorch, Tensorflow, Pandas, Torch Geometric, Deep Graph Library, Huggingface