JASON YANG

Pasadena, CA 91106 | (224) 255-9010 | jyang4@caltech.edu

OBJECTIVE: Dedicated graduate research student seeking opportunities to develop innovative, machine learning-assisted workflows for protein engineering.
Website: https://jsunn-y.github.io

EDUCATION

California Institute of Technology: Pasadena, CA; Chemical Engineering (Ph.D.) — (2021-2026)
Advisors: Professor Frances Arnold and Professor Yisong Yue
- o GPA: 4.0/4.0
- o **Teaching Assistant:** Thermodynamics, Transport

Yale University: New Haven, CT; Chemical Engineering (B.S.) & Applied Math (B.S.) — (2017-2021)
- o GPA: 3.93/4.00

New Trier High School: Winnetka, IL (2013-2017)
- o Weighted GPA: 5.33/5.33

AWARDS & HONORS
- o Caltech CCE Seminar Day Presentation Winner (2023)
- o National Science Foundation Graduate Research Fellowship (2021)
- o Barry Goldwater Scholarship (2020)
- o Tau Beta Pi Engineering Honor Society (2019)
- o First-Year Summer Research Fellowship, Yale University (2018)
- o Regeneron Science Talent Search Top Scholar, formerly Intel STS Semifinalist (2017)
- o National Merit Scholarship Finalist (2017)
- o Congressional Award for Youth Gold Medal, Washington DC (2017)
- o Science Olympiad National Medalist (2017)

RESEARCH & INDUSTRY EXPERIENCE

Summer Intern at Profluent, Berkeley, CA. Supervisors: Jeffrey Ruffolo and Aadyot Bhatnagar (Summer 2024)
- o Conditional generation of enzyme sequences using protein language models

Graduate Researcher at Caltech– Pasadena, CA. Supervisors: Professor Frances Arnold and Professor Yisong Yue (Fall 2021 – Present)
- o Developing machine learning-assisted workflows to accelerate protein engineering
  - o Representation learning using physics-informed features and graph neural networks
  - o Uncertainty quantification and Bayesian optimization with deep kernels
  - o Protein language models for enzyme generation
  - o Enzyme retrieval and annotation using multimodal models
- o Mentored two graduate rotation students and three undergraduate students

Undergraduate Research Assistant at Yale University– New Haven, CT. Supervisor: Professor Menachem Elimelech (Spring 2018 – Spring 2021)
- o Investigated cation-specific selectivity in membranes by synthesizing polyelectrolytes and incorporating them into novel nanofiltration membranes using layer-by-layer assembly
- o Studied the mechanism of ecotoxicology and environmental impact of MnO_2 nanomaterials using dynamic light scattering and experimental assays

NSF Research Experience for Undergraduates at the University of Connecticut– Storrs, CT. Supervisor: Professor Ying Li (Fall 2019 – Spring 2021)
- o Evaluated the parameters controlling water-organic separations in nanoporous graphene membranes using molecular dynamics simulations

DoE Science Undergraduate Laboratory Intern at the National Renewable Energy Lab– *BioEnergy Science & Technology*, Golden, CO. Supervisor: Dr. Heather Mayes (Summer 2020)

- o Implemented machine learning models using features from molecular dynamics simulations to accurately predict mechanical properties of bio-based polymers

Pharma Technical Development Intern at Genentech– *Analytical Operations*, South San Francisco, CA. Supervisor: <u>Steve Russell</u> (Summer 2019)
- o Automated an ICP-MS trace-metals assay for greater throughput by developing software and hardware

Summer Research Intern at Northwestern University– Evanston, IL. Supervisor: <u>Professor Teri Odom</u> (Summer 2016)
- o Developed a cost-effective nanoscale patterning method of polyolefin thin films for enhanced hydrophobicity

PUBLICATIONS

- o **Yang, J.** et al. Benchmarks for the Classification and Retrieval of Enzymes (in preparation)
- o **Yang, J.**; Lal, R. et al. Active Learning-Assisted Directed Evolution (in preparation)
- o Johnston, K.E.; Watkins-Dulaney, E.L.; Almhjell, P.J.; Liu, G.; Porter, N.J.; **Yang, J.**; Arnold, F. H. A combinatorially complete epistatic fitness landscape in an enzyme active site. (under review)
- o **Yang, J.**; Li, F.-Z.; Arnold, F. H. Opportunities and Challenges for Machine Learning-Assisted Enzyme Engineering. *ACS Cent. Sci.* **2024**, acscentsci.3c01275. <u>https://doi.org/10.1021/acscentsci.3c01275</u>.
- o **Yang, J.**; Ducharme, J.; Johnston, K. E.; Li, F.-Z.; Yue, Y.; Arnold, F. H. DeCOIL: Optimization of Degenerate Codon Libraries for Machine Learning-Assisted Protein Engineering. *ACS Synth. Biol.* **2023**, acssynbio.3c00301. <u>https://doi.org/10.1021/acssynbio.3c00301</u>.
- o **Yang, J.**; Tao, L.; He, J.; McCutcheon, J. R.; Li, Y. Machine Learning Enables Interpretable Discovery of Innovative Polymers for Gas Separation Membranes. *Sci. Adv.* **2022**, *8* (29), eabn9545. <u>https://doi.org/10.1126/sciadv.abn9545</u>.
- o He, J.; **Yang, J.**; McCutcheon, J. R.; Li, Y. Molecular Insights into the Structure-Property Relationships of 3D Printed Polyamide Reverse-Osmosis Membrane for Desalination. *Journal of Membrane Science* **2022**, *658*, 120731. <u>https://doi.org/10.1016/j.memsci.2022.120731</u>.
- o DuChanois, R. M.; Heiranian, M.; **Yang, J.**; Porter, C. J.; Li, Q.; Zhang, X.; Verduzco, R.; Elimelech, M. Designing Polymeric Membranes with Coordination Chemistry for High-Precision Ion Separations. *Sci. Adv.* **2022**, *8* (9), eabm9436. <u>https://doi.org/10.1126/sciadv.abm9436</u>.
- o **Yang, J.**; Shen, Z.; He, J.; Li, Y. Efficient Separation of Small Organic Contaminants in Water Using Functionalized Nanoporous Graphene Membranes: Insights from Molecular Dynamics Simulations. *Journal of Membrane Science* **2021**, *630*, 119331. <u>https://doi.org/10.1016/j.memsci.2021.119331</u>.
- o Zucker, I.; Hashmi, S. M.; **Yang, J.**; He, Y.; Pfefferle, L. D.; Elimelech, M. Shape-Dependent Interactions of Manganese Oxide Nanomaterials with Lipid Bilayer Vesicles. *Langmuir* **2019**, *35* (43), 13958–13966. <u>https://doi.org/10.1021/acs.langmuir.9b02428</u>.

COURSEWORK

California Institute of Technology (<u>Ph.D.</u>)
- o **Chemical Engineering Coursework:** Kinetics, Thermodynamics, Statistical Mechanics, Transport
- o **Other Coursework:** Machine Learning, Advanced Topics in Machine Learning, Deep Learning for Biology, Bioinformatics, Biomolecular Engineering, Enzymology and Biochemistry, Computational Tools for Metagenomics

Yale University (<u>B.S.</u>)

- **Chemical Engineering Coursework:** Organic Chemistry, Physical Chemistry, Materials Science, Programming for Engineers, Fluid Mechanics, Thermodynamics, Transport Phenomena, Chemical Kinetics and Reactors, Separations and Purifications
- **Applied Math Coursework:** Vector Calculus, Linear Algebra, Differential Equations, Data Analysis, Probability Theory, Applied Numerical Methods for Differential Equations, Statistical Methods for Science, Machine Learning, Discrete Math
- **Elective Courses:** Environmental Physiochemical Processes, Computational Chemistry, Inorganic Chemistry, Solid State Physics, Biochemistry and Cell Biology
- **Other Courses:** Tokyo, Fiction and Consciousness, Graphic Design, Intro to Psychology, Healthcare Economics and Public Policy, European Intellectual History, Advanced Culture and Conversation in French

SKILLS

Machine Learning Architectures: transformers, graph NNs, diffusion models, deep kernels, contrastive learning, self-supervision, VAEs, GANs, CNNs, RNNs
Python: Pytorch, Tensorflow, GPytorch, Botorch, Scikit, xgboost, Numpy, Pandas, Seaborn, Matplotlib, Bokeh, Holoviews, Biopython
Other Computer Software: Rosetta, Pymol, AlphaFold, BLAST, MUSCLE, HMMER, Autodock, LAMMPS, OpenMM, Gaussian, CHARMM, Materials Studio, CAD, Mathematica, C/C++, R, Adobe Suite
Lab Skills: Methods in directed evolution of proteins, HPLC, TLC, TOC, ICP-MS, ion chromatography, FTIR, UV-VIS, NMR, SEM, AFM, dynamic light scattering, spin coating, ellipsometry, organic synthesis, nanoscale surface modification of polymer materials
Other: graphic design, piano, advanced Mandarin Chinese and French, Chinese calligraphy