

## Kevin Yang

yangkky@gmail.com | (614) 260-9454 | <https://yangkky.github.io/about/>

### EDUCATION

**California Institute of Technology**, Pasadena, California (August 2014 – present)

*PhD candidate in Chemical Engineering. Expected 2018 graduation.*

Relevant coursework: Advanced Topics in Machine Learning; Introduction to Biomolecular Engineering; Machine Learning and Data Mining; Linear Algebra; Bioinformatics; Data Analysis for the Biological Sciences; Complex Analysis

**The Ohio State University**, Columbus, Ohio (Sept 2007 – June 2011)

*B.S. in Chemical and Biomolecular Engineering; minor in Music*

### EXPERIENCE

**Graduate Research Assistant**, Professor Frances Arnold's Group

*California Institute of Technology, Pasadena, CA (August 2014 - Present)*

- Used Gaussian process models ([github.com/yangkky/gpmodel](https://github.com/yangkky/gpmodel)) and design channelrhodopsins with improved properties
- Designed embedded representations of protein sequences based on doc2vec to streamline machine learning pipelines ([github.com/fha\\_lab/embeddings\\_reproduction](https://github.com/fha_lab/embeddings_reproduction))
- Built neural machine translation models on PyTorch to predict signal peptides from their corresponding mature protein sequences

### **Computational Intern**

*Ambry Genetics, Aliso Viejo, CA (June 2017 – September 2017)*

- Developed and implemented neural network models in Keras and PyTorch to predict outcomes of genetic variation by transferring information across paralogous proteins
- Incorporated model into a pipeline that finds paralogs for variants of interest and then uses paralogs and model to predict variant outcomes

### SELECTED PUBLICATIONS AND PRESENTATIONS

- **Yang, K. K.**, Wu, Z., Bedbrook, C. N., Arnold, F.H. "Protein embeddings for machine learning." *Bioinformatics*. 23 March 2018. <https://doi.org/10.1093/bioinformatics/bty178>
- Bedbrook, C. N., **Yang, K. K.**, Rice, A. J., Gradinaru, V., Arnold, F.H. "Machine learning to design integral membrane channelrhodopsins for efficient eukaryotic expression and plasma membrane localization". *PLOS Comp. Bio.* 23 Oct 2017. [doi.org/10.1371/journal.pcbi.1005786](https://doi.org/10.1371/journal.pcbi.1005786)
- Bedbrook, C. N., Rice, A. J., **Yang, K. K.**, Ding, X., Chen, S., LeProust, E. M., Gradinaru, V., Arnold, F.H. Structure-guided SCHEMA recombination generates diverse chimeric channelrhodopsins. *PNAS*. 10 Mar 2017. [doi/10.1073/pnas.170026911](https://doi.org/10.1073/pnas.170026911)
- **Oral Presentation:** "Machine Learning to Predict Eukaryotic Expression and Plasma Membrane Localization of an Integral Membrane Protein." Proteins Gordon Research Seminar, Holderness, NH. 17 June 2017.