# Specific Aims:

Alternative splicing is of fundamental importance to gene regulation, but the mechanisms behind it are incompletely understood. An interesting but poorly understood case of alternative splicing human mutually exclusive exons. Understanding the mechanisms that create mutually exclusive exons will both improve our understanding of fundamental processes and guide development of splice modifying therapies for human diseases. Massively parallel splicing assays (MPSAs) provide a promising method for understanding these mechanisms, but have technical limitations because they rely on short read sequencing. mutually exclusive exons are difficult to study with current MPSAs because the correlated decisions occur over distances longer than short reads.

I am interested in massively parallel experimental and computational methods to ask biological questions. In my doctorate, I applied these lenses to microbiology, protein engineering, and CRISPR tool development. Splicing is perfectly suited to this approach, with complex regulation amenable to MPSAs and quantitative modeling. By applying my skills to splicing, I will learn new applications of these techniques in a field with deep fundamental and applied questions. This will provide skills I need to begin my own independent quantitative biology lab.

I propose to develop a long read based MPSA (LR-MPSA), to create LR-MPSA analysis software, and apply these to dissecting the mechanisms of mutually exclusive exon splicing in the specific context of pyruvate kinase M (PKM). I will accomplish this in three complementary aims.

**Aim 1: Develop a nucleotide-resolution MPSA using long read sequencing***.* I will develop an LR-MPSA optimized for characterizing the mechanisms of complex splicing decisions in the context of full native introns and all isoform outputs. I will adapt MPSA techniques honed by the Kinney lab by creating new sample and library preparation pipelines optimized for unbiased isoform enrichment and nanopore sequencing.

**Aim 2: Develop software for analyzing data from nucleotide-resolution MPSAs***.* Calling low abundance RNA isoforms from nanopore reads faces challenges from the low depth and high noise. I will develop analysis software for identifying and quantifying isoforms that optimizes for power to detect low abundance isoforms and minimizes incorrect isoform assignment. I will simulate nanopore reads given a known ground truth isoform distribution for each variant to test and optimize the software at various depths and error profiles. I will make this software publicly available and open-source.

# Aim 3: Define the mechanism of mutually exclusive exon inclusion in PKM. PKM has a pair of mutually exclusive exons and switching between them drives the Warburg effect, this splicing event is a target for cancer therapies. PKM mutually exclusive exon splicing causes a shift from respiration to fermentation that is a driver of cancer. I will use LR-MPSAs to identify motifs that are important for maintaining mutual exclusivity in PKM mutually exclusive exons by identifying mutants that deviate from mutually exclusive exon behavior. I will verify identified motifs with anti-sense oligo based blocking and low throughput experiments, then model their behavior.

Concurrent with this research, I propose a directed effort to gain training and mentorship for my future career as an independent scientist. I will attend CSHL meetings and Gordon conferences on RNA processing and quantitative methods, attend lab meetings and journal clubs in the Kinney, Krainer, Koo and McCandlish labs, attend CSHL grant writing and professional development courses, and hone my skill at lecturing through teaching opportunities.

Together, these research and training opportunities will position me to launch an independent research career focusing on applying massively parallel assays and modeling to understanding deep biological questions in RNA processing.