Title: New Chemical Approaches for Studying Glycans and Proteins with Single-Cell Resolution

Authors: Katsuya L. Colón,* Terry D. Kim,* John W. Thompson,* Jean-Luc Chaubard, Chithra Krishnamurthy, Wen Yi, Heather E. Murrey, Cristal I. Gama, Stacey A. Kalovidouris, Wen.-I. Luo, and Linda C. Hsieh-Wilson.

*These authors contributed equally.

Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, CA 91125.

Abstract: Gross disruptions in protein glycosylation have been universally identified as a feature of Alzheimer's disease (AD), but how these disruptions contribute to the disease process is largely unknown. Current methods for studying glycans focus on identifying changes in glycan expression across large populations of cells, tissues, and/or disease states. However, these populations often have a diverse cellular make-up, and the same glycans may have vastly different roles in different cell types or contexts. Moreover, glycans are often analyzed separately from the proteins they are attached to or using non-specific detection techniques. To date, there are no methods capable of providing an unbiased, systems-level assessment of glycan expression levels, for tracking how they vary with cell type or state, or for detecting and quantifying defined glycoproteins across the proteome. In response to these challenges, we are working to develop Single-cell Unified Glycan and RNA-sequencing (SUGAR-seq), a new method that seamlessly integrates chemoenzymatic labeling of specific glycans with multimodal single-cell RNA sequencing to quantify specific glycan expression levels simultaneously with the transcriptome at the level of single cells. This technology will allow us to determine for the first time: (1) which cell types display certain glycans and at what levels, (2) how changes in glycan expression are associated with changes in the transcriptome, and (3) which glycoprotein biosynthesis, transport, and processing genes are differentially expressed across different cell types and states in both healthy and diseased brains. To also understand how underlying glycoprotein dynamics contribute to changes in glycan expression, we will use the same chemoenzymatic labeling techniques to enrich glycoproteins containing glycans of interest for analysis by quantitative mass spectroscopy. Importantly, this will also reveal specific glycoproteins or glycosylation sites that may be altered by cell or disease state for targeted functional studies. Overall, this novel multi-omics approach will allow us to interrogate the expression and dynamics of specific glycans and glycoproteins in neurophysiology and AD.