UNIVERSITY OF CALIFORNIA

Los Angeles

Astrocyte and neuron subproteomes reveal cell-specific contributions relevant to obsessive-compulsive disorder behaviors

A dissertation submitted in partial satisfaction of the
requirements for the degree Doctor of Philosophy
in Neuroscience

by

Joselyn Sarai Soto

2023

ABSTRACT OF THE DISSERTATION

Astrocyte and neuron subproteomes reveal
cell-specific contributions relevant to obsessive-compulsive disorder behaviors

by

Joselyn Sarai Soto
Doctor of Philosophy in Neuroscience
University of California, Los Angeles, 2023
Professor Baljit S. Khakh, Chair

Astrocytes are morphologically complex and bushy cells that tile the entire central nervous system and contribute, along with neurons, to the formation of a complex network with dense physical, chemical, and electrical connectivity. In the striatum, astrocyte fine processes contact about eleven neuronal somata and their synapses where they mediate homeostatic and supportive roles. Because there is a clear spatial and functional relationship between astrocytes and neurons, we used cell- and subcompartment-specific proximity dependent biotinylation to explore the proteomes of striatal astrocytes and neurons *in vivo*. We studied cytosolic and plasma membrane compartments of astrocytes and neurons to assess how these cells differ at the protein level. Furthermore, we assessed functional subcellular compartments of astrocytes, including the end feet and fine processes, to understand the molecular basis of essential astrocytic functions. Surprisingly, SAPAP3 protein (gene: *Dlgap3*) associated with obsessive-compulsive disorder (OCD) was detected at equivalent levels in both cell types and in select astrocyte subcompartments

including the plasma membrane and fine processes. Validation with transcriptomics, *in situ* hybridization, and protein labeling revealed that SAPAP3 RNA and protein are highly expressed in astrocytes. Furthermore, causal genetic experiments in the SAPAP3 knockout mouse model of OCD revealed cell-specific contributions of astrocytes and neurons to repetitive and anxiety-related behaviors. Thus, our data define the molecular basis of astrocyte and neuronal signaling in the striatum and how astrocyte and neuronal SAPAP3 mechanisms contribute to OCD phenotypes in mice. Taken together, our findings suggest that targeting both astrocytes and neurons could likely be effective in OCD and potentially other neuropsychiatric disorders.

Supplemental Materials (attached separately)

Supplemental movie 1: Astrocyte subproteome cards

Supplemental movie 2: SAPAP3 KO mice treated with either astrocyte-selective GFP or astrocyte-selective GFP-SAPAP3 are shown grooming

Supplemental movie 3: SAPAP3 KO mice treated with either neuron-selective GFP or neuron-selective GFP-SAPAP3 are shown grooming

Supplemental Excel File 1: Striatal astrocyte and neuron proteomes

Table 1: Astrocyte cytosol proteins

Table 2: Astrocyte plasma membrane proteins

Table 3: Astrocyte SAPAP3 interactors

Table 4: Neuron cytosol proteins

Table 5: Neuron plasma membrane proteins

Table 6: Neuron SAPAP3 interactors

Supplemental Excel File 2: Astrocyte subproteomes

Supplemental Excel File 3: RNA-sequencing data from astrocytes and neurons in the striatum expressing cell-selective RiboTag (Rpl22-HA)

The dissertation of Joselyn Sarai Soto is approved.

Kate Wassum

James A. Wohlschelgel

Stephen Lawrence Zipursky

Baljit S. Khakh, Committee Chair

University of California, Los Angeles

2023

For my father, Cesar H. Soto

For my mother, H. Elizabeth Soto

For my grandmother, Gregoria Serrano

Con todo mi amor.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ACRONYMS

AAV: adeno-associated virus

ATP: adenosine 5’ triphosphate

Ca^{2+}: calcium ions

Co-IP: co-immunoprecipitated

DNA: deoxyribonucleic acid

DCZ: deschloroclozapine

DE: differentially expressed

DREADD: designer receptor exclusively activated by designer drug

FACS: fluorescence-activated cell sorting

FC: fold-change

FDR: false discovery rate

FISH: fluorescent *in situ* hybridization

FPKM: fragments per kilobase of transcript per million mapped reads

GFP: green fluorescent protein

GO: gene ontology

GPCR: G-protein coupled receptor

HEK: human embryonic kidney

IHC: immunohistochemistry

IP: immunoprecipitated

ISH: *in situ* hybridization

K^{+}: potassium ions

KO: knock-out

LC-MS/MS: liquid chromatography-coupled tandem mass spectrometry

MSN: medium spiny neuron

mRNA: messenger ribonucleic acid

OCD: obsessive-compulsive disorder

P: post-natal day

PANTHER: protein analysis through evolutionary relationships

PCA: principal component analysis

PLA: proximity ligation assay

PM: plasma membrane

PSD: postsynaptic density

ROI: region of interest

RNA-Seq: ribonucleic acid sequencing

SAINT: significance analysis of interactome

tdT: tdTomato red fluorescent protein

WT: wild-type

ACKNOWLEDGEMENTS

This doctoral dissertation was completed because of the many people who contributed to the science and my growth as a scientist. First and foremost, I would like to express my heartfelt gratitude to my mentor, Dr. Baljit Khakh, for giving me the opportunity to undertake this project and for his unwavering support and patience these last four years. Thank you for catalyzing my career as a scientist. I would also like to express thanks to my committee members, Dr. Kate Wassum, Dr. James Wohlschlegel, and Dr. Larry Zipursky for their time and feedback to my work. I would especially like to thank Dr. James Wohlschlegel for his collaboration to make the proteomics section of this dissertation possible. His valuable advice helped move the mass spectrometry experiments forward. Many thanks to the present and past members of the Khakh lab for technical training and advice throughout the course of my work. My training was made possible due to financial support from the UCLA Department of Physiology, the UCLA Neuroscience Interdepartmental Program, the UCLA Graduate Division Cota-Robles Fellowship, and the National Science Foundation Graduate Research Fellowship Program (DGE-2034835). The project was supported by Dr. Khakh's grants from the National Institutes of Health (R35 NS111583), an Allen Distinguished Investigator Award, and the Ressler Family Foundation.

The work described throughout this dissertation, with the exception of Chapter 7, is modified from an accepted manuscript in press: Joselyn S. Soto, Yasaman Jami-Alahmadi, Jakelyn Chacon, Stefanie L. Moye, Blanca Diaz-Castro, James A. Wohlschlegel, and Baljit S. Khakh. Astrocyte-neuron subproteomes and obsessive-compulsive disorder mechanisms. Nature (2023). The entirety of the proteomic work described in **Chapters 3 and 4**, was done in close collaboration with Dr. Yasi Jami-Alahmadi from Dr. Wohlschlegel's lab. Specifically, Dr. Jami-Alahmadi operated the mass spectrometer and helped develop the analysis pipeline by which we compared the astrocyte

and neuron proteomes and the astrocyte subproteomes. The neuronal and astrocytic BioID2 plasmid constructs were cloned by Jacky Chacon under the guidance of Dr. Blanca Diaz-Castro. Data for **Figures 3.3 and 3.4** were generated by Jacky Chacon under the guidance of Dr. Blanca Diaz-Castro. Dye filling of complex astrocytes in **Figure 1.2** was conducted by Stefi Moye. RNA-sequencing was performed at the UCLA Neuroscience Genomics Core. Dr. Fuying Gao helped with the analysis of RNA-sequencing data. Dr. Vijaya Pandey helped by performing mass spectrometry for **Figure 5.12**. **Chapter 6** was made possible with contributions from Dr. Xinzhu Yu. With her help, we were able to establish the SAPAP3 KO mouse line and replicate the OCD-like behaviors. Stefi Moye helped with **Figures 6.1 and 6.2** by conducting mouse behavioral analyses under the guidance of Dr. Xinzhu Yu.

Thank you to Dr. Fumito Endo for help and advice with immunohistochemistry, and single-cell sequencing analysis. Thank you to Dr. Ling Wu for help with AAV injections conducted for **Chapter 4**. Thanks to Drs. Avishek Adhikari and My La-Vu for helpful advice on mouse anxiety behaviors and analysis. To Dr. Guoping Feng for the gift of the SAPAP3 antibody. To Dr. Scott Soderling for gifting the BirA plasmid. Thanks to Dr. Tom Vondriska for use of the sonicator. Many thanks to Khakh lab members who provided intellectual engagement and friendly support.

My PhD training has been incredibly rewarding and challenging. Thus, I would like to thank Dr. Baljit Khakh, Khakh lab members, the Department of Physiology administration, and members of the UCLA Neuroscience program for their encouragement and moral support throughout my journey here. Thank you to Kevin Martinez for all his loving support. My sisters, Jackie and Stefanie for motivating me and encouraging my dreams. Finally, thank you to my father and mother. Had it not been for their decision to emigrate from Guatemala to the United States to give my sisters and me a better life, this would not have been possible.

VITA

Education

2012-2017 B.S. in Cell, Molecular, and Developmental Biology **University of California, Riverside,** Riverside, CA, USA

Publications

1. **Soto JS**, Jami-Alahmadi Y, Chacon J, Moye SL, Diaz-Castro B, Wohlschlegel JA, and Khakh BS (2023). Astrocyte-neuron subproteomes and obsessive-compulsive disorder mechanisms. ***Nature*** (Article) 616, 764-773 (2023). doi: https://doi.org/10.1038/s41586-023-05927-7.

2. Ollivier M, Linker K, **Soto JS**, Moye SL, Jami-Alahmadi Y, Jones A, Divakaruni A, Kawaguchi R, Wohlschlegel J, and Khakh BS (2023) Subregionally allocated u-crystallin expressing astrocytes regulate perseveration (*in preparation*)

3. Gangwani M, **Soto JS**, Jami-Alahmadi Y, Tiwari S, Kawaguchi R, Wohlschlegel J, and Khakh BS (2023). Neuronal and astrocytic contributions to Huntington's disease dissected with zinc finger protein transcriptional repressors. ***Cell Reports*** (Article) 42, 111953. doi: 10.1016/j.celrep.2022.111953. PMID: 36640336

4. Endo F. Kasai A, **Soto JS**, Yu X, Qu Z, Hashimoto H, Gradinaru V, Kawaguchi R, and Khakh BS (2022). Molecular basis of astrocyte diversity and morphology across the CNS in health and disease. ***Science*** (Research Article) Nov 4; 378(6619). doi: 10.1126/science.adc9020. PMID: 36378959

5. Yu X, Nagai J, Marti-Solano M, **Soto JS**, Coppola G, Babu MM, and Khakh BS. (2020) Context-specific striatal astrocyte molecular responses are phenotypically exploitable. ***Neuron*** 108: 1146-62. doi: 10.1016/j.neuron.2020.09.021. PMID: 33086039

6. Sekyi M, Lauderdale K, Atkinson K, Gokestany B, Karim H, Feri M, **Soto JS**, Diaz C, Kim S, Cilluffo M, Nusinowitz S, Katzenellenbogen J, and Tiwari-Woodruff S. (2021) Alleviation of extensive visual pathway dysfunction by a remyelinating drug in a chronic mouse model of multiple sclerosis**.** ***Brain Pathology*** 130(2). doi: 10.1111/bpa.12930. PMID: 33368801

7. Atkinson K, Lee J, Hasselmann J, Kim S, Drew A, **Soto JS**, Katzenellenbogen J, Harris N, Obenaus A, and Tiwari-Woodruff S. (2019) Diffusion tensor imaging identifies aspects of therapeutic estrogen receptor β ligand-induced remyelination in a mouse model of multiple sclerosis. ***Neurobiology of Disease*** (130). doi: 10.1016/j.nbd.2019.104501. PMID: 31226301

Honors and Awards

2023	UCLA Life Science Excellence Award: Excellence in Research
2023	31st Annual Samuel Eiduson Student Lecture award at UCLA
2022	LABEST 2022 Pearl Cohen Poster Award (2nd place)
2017-2020	National Science Foundation Graduate Research Fellowship Program
2017-2020	UC Eugene V. Cota-Robles Fellowship
2017	Fulbright Scholarship: Germany
2015-2017	National Institutes of Health MARC-U Trainee
2016	SACNAS Travel Award Recipient
2016	SACNAS Outstanding Poster Award

CHAPTER 1: INTRODUCTION

Section 1.1: Historical perspective on astrocytes

While giving a lecture on a warm spring day in 1858, famed physician, Rudolf Virchow announced the concept of 'neuroglia,' or nerve glue, to a class of medical students. Virchow understood that neurons, the electrical conduit cells of the nervous system, did not exist alone in the nervous system. However, Virchow deduced that this neuroglia was a connective-like tissue that merely provided structural support to the neurons and did not contain any cells (Virchow, 1856; Virchow, 1858).

The first conclusive investigation of neuroglia came from Camillo Golgi in the early 1870s when he developed a robust method to stain cells with silver nitrate solution. Golgi's method, as this staining procedure is now called, allowed visualization of single cells and in this way the cells making up the neuroglia were visualized. Golgi was able to visualize and illustrate the first images of glial-vascular contacts (Golgi, 1873; Golgi, 1903) and pioneered the first theory of glial function which proposed that glial cells of the central nervous system were responsible for mediating substance exchange between the vasculature, thus providing metabolic support to the rest of the brain through these specialized glial-vascular contacts. Because Golgi's intricate illustrations showed these glial cells as stellate with protruding branches, Hungarian anatomist, Michael von Lenhossek named these cells 'astrocytes' (Lenhossek, 1891).

As the 19th century came to a close, novel anatomical methods began to emerge and astrocyte research advanced as Santiago Ramon y Cajal developed a gold sublimate staining method that allowed visualization of astrocytes (Ramon y Cajal, 1913). With this method, Cajal found that astrocytes and neurons spatially and intimately interact and thus, he thought that astrocytes perhaps worked spatially to regulate neuronal information flow (Parpura and Verkhratsky, 2012). As visualizations and illustrations of astrocytes began to emerge in this time period, the idea that

astrocytes and neurons functionally interact began to take shape. German surgeon, Carl Ludwig Schleich, proposed that astrocytes could be driving general inhibitory mechanisms in the brain via their peri-synaptic processes and that the size of these processes could control information flow through synapses (Schleich, 1894). Cajal also suggested that the swelling of peri-synaptic processes could block neurotransmission and was a mechanism of sleep (Ramon y Cajal 1925).

Finally, astrocytes began to become implicated in pathological contexts. Alzheimer began to contemplate the role of astrocytes in disease, including dementia (Alzheimer 1910), while Andriezen observed that in pathological conditions, astrocytes became 'hypertrophic' (Andriezen, 1893). While these observations led to exciting questions emerging in the novel astrocyte field, the function of astrocytes remained unexplored for the next part of the century. As novel methods for recording and studying the electrically excitable neurons began to emerge, research on the electrically "silent" astrocytes was neglected. Furthermore, because of the technical advents of electrophysiology, the neuroscience field began to adopt a neuro-centric doctrine in which neurons are considered the sole drivers of brain signaling, disease, and behavior. As the novelty of single-cell electrophysiology decreased, the past forty years have seen a plethora of data supporting the early ideas of astrocytes as functional components of the central nervous system and its function. However, outstanding questions remain in our understanding of astrocyte function and its relation to neuronal function.

Section 1.2: Astrocytes play essential roles in the central nervous system

In order for the central nervous system to function properly, adequate maintenance of the microenvironment is essential for normal neuronal function. This includes regulation of the extracellular pH, osmolarity, extracellular ion composition, and the neurotransmitter concentration

(Benarroch, 2005). Furthermore, as deduced by Golgi, there should be an abundant supply of fuel for neuronal metabolism and local regulation of blood flow in order to meet the energy demands of active populations of neurons.

Astrocytes are the most abundant glial cells in the brain and make up about half of the cells in the human adult brain (Barres, 2008). Using specialized genetic and imaging tools for astrocytes, we now know that astrocytes are remarkably complex and "bushy" cells with thousands of finger-like processes extending from the cell body (Bushong et al., 2002; Shigetomi et al., 2013; Moye et al., 2019). The finest of these processes interact intimately with synapses where they can regulate neurotransmitter concentration, synapse formation, synapse removal, and overall synapse function (Allen, 2014; Allen and Lyons, 2018). Furthermore, astrocytes contain a specialized perivascular endfoot which fully ensheathes the cerebral vasculature, regulates blood flow, regulates nutrient uptake and contributes to the blood brain barrier (Nuriya and Yasui, 2013; Belanger et al., 2011; Bazargani and Attwell, 2016). Additionally, astrocytes are the only cells in the central nervous system that store glycogen (Ventura and Harris 1999; Lundgaard et al., 2014) and are a major source of cholesterol and lipoproteins, which are structural components of cell membranes.

As Cajal and Schleich had reasoned, recent evidence has also shown that astrocytes can influence neuronal function. Neurons and astrocytes are intimately associated and form two separate interactive networks: the neuronal network which is connected by synapses and an astrocyte network connected via gap junctions composed of connexin proteins to form a syncytium. These interactions allow astrocytes to modulate neuronal function by regulating neuronal excitability. Indeed, astrocytes facilitate repetitive neurotransmission through uptake of potassium ions (K+) and neurotransmitters such as glutamate and GABA (Walz 2000, Oliet et al., 2001). Because astrocytes are highly permeable to K+ ions, and thus, their very low resting

potential does not allow neuronal-like action potentials to occur, these cells have been canonically considered as electrically "silent" and unresponsive. This is mostly because the electrophysiological tools by which neurons are studied cannot be used to the same effect in astrocytes. However, the advent of genetically targeted and cell-specific tools have demonstrated that astrocyte responses exist via changes in intracellular calcium (Ca^{2+}). Ca^{2+} changes have been observed in astrocytes in response to local neuronal activity (Charles et al., 1991; Dani et al., 1992; Kang et al., 1998) and during discrete behavioral responses *in vivo* such as self-grooming and locomotion in mice (Paukert et al., 2014, Shigetomi et al., 2016, Srinivasan et al., 2015, Yu et al., 2018). Evidence also shows that these astrocyte Ca^{2+} signals are directly involved in complex mouse behaviors such as self-grooming (Yu et al., 2018; Halassa and Haydon 2010).

While we have learned much about astrocyte function in the brain over the last forty years with the advents of new tools catered for astrocytes, the fundamental biological mechanisms by which astrocytes and neurons interact *in vivo* are still unknown. Furthermore, the channels and molecules implicated in regulating Ca^{2+} pathways, regulating endfoot function, and regulating peri-synaptic processes have not been fully described. This missing puzzle piece would give great insight on how astrocytes fulfill their many functions in the central nervous system.

Section 1.3: Molecular methods to study astrocytes

With the genomic revolution occurring at the beginning of the 21st century and molecular methods such as genomic sequencing and RNA-sequencing becoming high-throughput and readily available, understanding the genes expressed by astrocytes in normal and pathological contexts in mouse models became crucial to understand basic astrocyte biology. Methods to isolate astrocytes in order to obtain cell-specific transcripts were developed including: primary cell culture (Bachoo

et al., 2004; Zhang et al., 2014), dissociation for fluorescence activated cell sorting (FACS) of genetically fluorescent astrocytes (Cahoy et al., 2008), cell-specific mouse lines expressing tagged ribosomes in astrocytes (Heiman et al., 2008; Srinivasan et al., 2016), and more recently, single cell sequencing (Yu et al., 2020). The advent of these methods allowed astrocyte RNA profiles to be studied extensively in a variety of contexts including throughout development / aging (Zhang et al., 2014; Clarke et al., 2018), during spinal cord injury (Anderson et al., 2016), in the context of stroke (Rakers et al., 2019), in different brain regions (Lin et al., 2017; Chai et al., 2017; Endo et al., 2022) and many more contexts (Yu et al., 2020). Cell-specific RNA sequencing methods have thus, allowed us to understand astrocytes in an unprecedented manner. For example, we now understand that not all astrocytes express glial acidic fibrillary protein (GFAP), the protein that Cajal's gold sublimate method stained. We now also understand that depending on the context, different molecular pathways within an astrocyte change which could give rise to a variety of functional responses.

While RNA sequencing has been invaluable in the quest to understand astrocyte function in health and disease, it is important to reassess what this readout signifies. While the central dogma of biology states that the nucleotide sequence of a gene determines the sequence of its mRNA product and this mRNA's sequence determines the amino acid sequence of a polypeptide chain, there is no easy way to explain the relationship between the concentration of an mRNA transcript and the concentration of the resulting protein (Liu et al., 2016). Post-translational modifications (Tang and Amon, 2013), translation rates (Wethmar et al., 2010), translational modifications (Barrett et al., 2012), and cell states (Liu et al., 2016) all have a complex influence on protein abundance and its relation to mRNA abundance. Furthermore, there are different types of mRNA-protein correlations that must be clearly identified (Vogel and Marcotte, 2012). For example, if an

astrocyte upregulates GFAP expression under a variety of conditions, the correlation of GFAP mRNA and protein could be measured across different conditions. This type of analysis would address to what extent the variation of protein levels is determined by variations of its corresponding mRNA levels. However, if multiple genes are changed under a certain condition, assessing the mRNA and protein level concentration of those genes would address to what extent concentration differences between transcripts of different genes show up at the protein level. (Vogel and Marcotte, 2012; Marguerat et al., 2012). Furthermore, because a cell in itself is not a stable entity, as it is constantly changing its energy needs and regulating itself, mRNA and protein levels can vary in relationship to each other both spatially and temporally. For example, if a new nascent polypeptide chain is being translated, the mRNA levels will be significantly higher than the protein levels at time = 0 versus at time = 30 sec. Therefore, it is important to consider how much we can interpret from mRNA in relation to astrocyte function.

Proteins are the main macromolecules that mediate function and are the primary output of gene expression. Proteins have discrete and precise engineered parts that are coupled to chemical events (Alberts et al., 2002) and are the only type of macromolecule that interacts with all other macromolecules such as lipids, carbohydrates, RNA, DNA, and other proteins to conduct their functions. Proteins are highly specialized to their interactors, are precisely localized within a cell, and have a vast repertoire of diverse functions. Therefore, a cell's functional output is highly dependent on the proteins it expresses. Recent advents in protein biology, including the inception of mass-spectrometry based proteomics forty years ago, have allowed analysis of proteins in diverse biological samples (Han et al., 2008). While a variety of different mass-spectrometry approaches have been developed over the last four decades, a generic workflow has emerged and has been continuously refined to ask address a variety of different questions in neuroscience. In

short, proteins from any source are extracted and prepared for enzymatic digestion. The resulting short peptides are separated based on selected parameters such as hydrophobicity or charge with liquid-chromatography. The liquid-chromatographer is coupled to a mass spectrometer via an electrospray ionization source which sprays the ionized peptides into the mass spectrometer. The mass and intensity of the peptides are then measured in a series of mass spectrometry scans that alternate between fragmentation and peptide selection so that peptides may be identified by comparing the fragments to a sequence database (Hosp and Mann, 2017). The use of this standard workflow allows the identification and quantification of the proteins present in the sample. In the protein field, this type of workflow is termed "bottom-up" and when coupled with up-front enrichment of a specific cell-type or a cell sub-compartment could identify the proteins found within a specific cell-type or compartment (Larance and Lamond, 2015). Therefore, the use of mass-spectrometry based proteomics is a valuable tool to assess the proteins expressed by different cell types, including astrocytes.

To obtain a readout on the proteins expressed by astrocytes, one must first be able to extract proteins only from astrocytes. As with RNA-sequencing, different methods have emerged to allow astrocyte-specific proteomes to be studied. Traditionally, astrocytes have been cultured (Yang et al., 2005; Han et al., 2014) or have been dissociated and isolated with FACS (Chai et al., 2017) in order to be processed for downstream mass-spectrometry proteomics analyses. However, under the microscope, both the cultured and the FACS isolated astrocytes look different from their *in vivo* and *in situ* counterparts. Therefore, it is important to acknowledge that the proteomes identified from these isolated astrocytes may not directly reflect the proteins of those that exist *in vivo* physiological conditions. Therefore, new tools are needed to reliably assess the protein profiles of astrocytes to assess their function in relation to neural circuitry.

Section 1.4: Astrocytes and neural circuits: focus on the striatum

In order to understand how astrocytes function within neural circuits, it is imperative to comprehensively study astrocytes in a single brain area to understand functions that are tailored to a specific and discrete circuit (Khakh, 2019). The basal ganglia (BG) consists of subcortical nuclei that are primarily responsible for motor control, motor learning, innate behaviors, executive behaviors, habit formation, and emotions (Lanciego et al., 2012; Graybiel and Grafton, 2015). The striatum is the largest input nucleus of the basal ganglia and is the largest subcortical brain structure in the mammalian brain with an estimated volume of 23 mm^3 in mice (Schröeder et al., 1975; Yin et al., 2009). The striatum receives excitatory input from various cortical and subcortical structures, such as the thalamus, and projects to several basal ganglia nuclei (Graybiel and Grafton, 2015). The neuronal populations in the striatum have been extensively characterized and have been shown to primarily consist of two populations: the spiny γ-aminobutyric acid (GABA) projection neurons termed medium spiny neurons (MSNs), which make up about 95% of total neurons, and interneurons, which make up about 5% of the total neuronal population and primarily serve to regulate local MSNs (Kawaguchi et al., 1995). MSNs consist of two populations which are determined by the type of dopamine receptor expressed by each: either D1-dopamine receptor expressing neurons or D2-dopamine receptor expressing neurons referred to as D1 MSNs or D2 MSNs, respectively. The primary function of striatal circuitry is to regulate which population of MSNs fire when excitatory input reaches the striatum which then influences the activity of downstream BG nuclei, and thus BG-mediated behaviors (Nicola et al., 2000; Grillner and Graybiel, 2006; Kreitzer and Malenka, 2008).

In brief, activation of D1 MSNs via the "direct" pathway inhibits the firing of the BG output nuclei including the substantia nigra pars reticulate (SNr) and globus pallidus (GPi), resulting in disinhibition of thalamic projections unto the cortex. However, activation of D2 MSNs via the "indirect" pathway enhances SNr activity which continues to chronically inhibit the thalamic projections unto the cortex. The complementary roles of D1 and D2 MSNs and their respective pathways, thus have important implications for pathophysiological disorders such as Parkinson's disease or Huntington's disease which produce disrupted motor functions such as dystonia, chorea, or bradykinesia (Lanciego et al., 2012; Bates et al., 2015). Furthermore, D1 and D2 pathways have been implicated in psychiatric disorders such as obsessive-compulsive disorder, attention-deficit hyperactivity disorders, habits, addictions, and tics (Graybiel, 2008; Burguiere et al., 2015). Because of its large size, its well characterized input / output, its well characterized MSN population, and its distinct anatomical units, the striatum presents itself as an ideal microcircuit in which to study astrocyte function in relation to neuronal circuitry.

Recent studies have begun to address outstanding questions in astrocyte biology using the striatum as a model circuit. In order to assess the identity of striatal astrocytes, molecular methods such as the ones detailed previously such as transcriptomics and were used. Transcriptomic analysis of adult mouse astrocytes using a mouse line with epitope-tagged ribosomes in astrocytes showed 1180 genes that were significantly enriched in striatal astrocytes when compared to hippocampal astrocytes, and 337 genes were significantly enriched in striatal astrocytes when compared to twelve other brain regions (Chai et al., 2017; Endo et al., 2022). There are also sets of genes that are both highly expressed and not differentially expressed in striatum which represent genes involved in core functions of astrocytes such as glutamate transporters and potassium channels. Together these data show that there are molecularly separable pathways in astrocytes

that are unique to striatum, but there are core functions that are shared among astrocytes across the central nervous system. When proteomics analysis of FACS isolated striatal astrocytes was conducted, 18 proteins were highly enriched in striatal astrocytes when compared to hippocampal astrocytes. One of these proteins was protein μ-crystallin (gene: *Crym*) which was also highly enriched in the transcriptomic datasets. The function of this protein in striatal astrocytes is unknown, nevertheless, the proteomics data showed unique striatal astrocyte functional identity in a similar manner to the transcriptomic data (Chai et al., 2017). A caveat of this study was that about 10% of the highest abundance proteins showed low mRNA abundance values, indicating the non-linearity of mRNA and protein abundance or as an artifact of FACS isolation not reflecting *in vivo* astrocyte physiology. To assess this, novel tools are needed to isolate striatal astrocyte proteins in a non-disruptive manner.

Functional studies of striatal astrocytes have also been conducted in order to assess their interactions with neurons. When astrocytes in the striatum are filled *in situ* with a diffusible fluorescent dye, the astrocyte looks "star-like" with a 20 μm diameter soma, about six primary branches arising from the soma, multiple smaller branches, one or more thick endfoot associating with blood vessels, and a cloud-like eminence comprised of thousands of fine finger-like processes (Chai et al., 2017, Moye et al., 2019). The entire volume that the astrocyte encompasses from its soma to its finest finger-like processes is called the astrocyte territory and in the striatum an average territory is about 60,000 μm^3 (Chai et al., 2017). Within a striatal astrocyte territory, there are an average of 12 NeuN positive neuronal somas interacting intimately with the astrocyte (**Figure 1.1**). Among the 12 neurons found within an astrocyte territory, there are about six D1

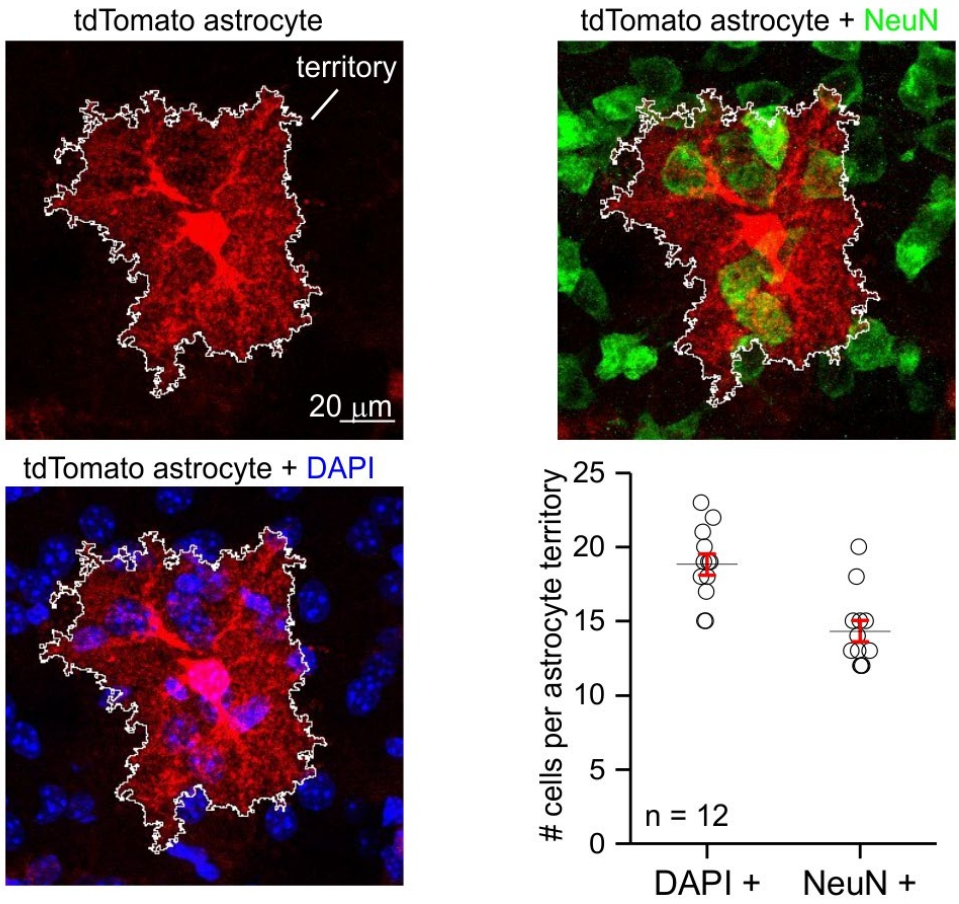

Figure 1.1: Astrocytes and neurons in the striatum spatially interact. Striatal astrocyte sparsely labeled with tdTomato fluorescent protein and co-stained with nuclear marker DAPI and neuronal marker NeuN. Quantification shows DAPI+ nuclei and NeuN+ cell bodies found in a single tdTomato+ astrocyte's territory (n = 12 tdT+ cells per group from 4 mice).

MSNs and five D2 MSNs which is similar to the relative densities of D1 and D2 MSNs within the striatum (Chai et al., 2017). Because striatal astrocytes are so proximal to many MSN somata, this suggests a spatially critical role of astrocytes in regulating the extracellular environment of MSNs. Due to the density of excitatory synapses in the striatum being about 0.9 / μm^3, this implies that an astrocyte territory can surround about 50,700 excitatory synapses (Ingham et al., 1998). These spatial interactions between striatal astrocyte processes with synapses were assessed with a fluorescence energy transfer (FRET)-based method and with high resolution serial block face scanning electron microscopy (SBF-SEM) (Chai et al., 2017; Octeau et al., 2018). The average distance from the pre or postsynaptic membrane to the membrane of the nearest astrocyte process was found to be about 50 nm, however, the distance between astrocyte processes and the pre-synaptic membrane was variable with the closest contacts being about 10 nm away, while the farthest were 400 nm away. This variability could indicate that striatal astrocytes make variable finger-like associations with synapses (Ventura and Harris, 1999; Khakh. 2019). Taken together, these structural data all indicate that astrocytes are spatially primed to conduct functions that regulate neuronal function.

Like neuronal Ca^{2+} signaling, astrocyte Ca^{2+} signaling is thought to be a major pathway by which astrocytes integrate and process information. Since the 1990s, it has been hypothesized that astrocyte Ca^{2+} signals may influence neuronal signaling and thus, neural circuits (Smith, 1994). Striatal astrocytes display a complex array of intracellular Ca^{2+} elevations in which Ca^{2+} enters through the plasma membrane and is also released from intracellular stores (Srinivasan et al., 2016; Jiang et al., 2016). In the striatum, attenuation of intracellular Ca^{2+} signaling by overexpression of a plasma membrane pump in mice showed an increase in excessive self-grooming behaviors (Yu et al., 2018). Molecular data in these mice showed that ambient GABA was reduced and thus,

astrocytes regulate striatal MSNs and concordant behaviors by GABA-mediated modulation of large volumes of tissue (Khakh 2019). Furthermore, when specific Ca^{2+} signaling pathways are activated in astrocytes, this too can cause changes in MSN physiology. When Gi-coupled G-protein-coupled receptor (GPCR) activation is conducted in striatal astrocytes, these show elevations in Ca^{2+} levels (Chai et al., 2017; Nagai et al., 2019). Release of GABA from MSNs activates endogenous GABA sensitive GPCRs in astrocytes and this activation of the endogenous Gi pathway results in elevation of intracellular Ca^{2+} levels in astrocytes. When chemogenetics were used to activate the Gi pathway even more, this resulted in elevated excitatory transmission, increased excitatory synapse formation, and increased MSN firing which led to a hyperactive phenotype with disrupted attention in mice (Nagai et al., 2019). Taken together, these data show that striatal astrocyte Ca^{2+} changes could lead to direct MSN modulation and changes in striatal linked behaviors.

While this past decade has indeed shown that astrocytes can play a causal role in complex behaviors, the precise mechanisms by which these roles are played remain unknown. Furthermore, the relative contributions of astrocytes and neurons to complex behaviors remain to be explored. Molecular methods for studying astrocytes have proven useful for a variety of contexts and brain regions, however at the heart of basic astrocyte biology lies the complex morphological structure that gives astrocytes their functional identity. The next step to understand astrocyte biology is to assess what proteins are expressed at the plasma membrane and other functionally relevant subcompartments in order to understand how astrocytes conduct their diverse functions at interfaces of cell-cell contacts in a well described brain circuit such as the striatum.

Section 1.5: Cell- and subcompartment-specific proteomes

In order to understand the basic biology of astrocytes and neurons, including how they interact, there is a necessity to assess protein identities and their fundamental differences *in vivo*. As mentioned previously, cell culture and cell dissociation procedures disrupt astrocyte and neuronal morphology (**Figure 1.2**). Furthermore, cultured astrocytes have been shown to be molecularly distinct from freshly dissociated astrocytes (Cahoy et al., 2008) indicating that the two may be molecularly dissimilar to intact astrocytes. Thus, the *in vivo* proteomes of astrocytes and neurons have not been directly measured or compared to understand their basic biology.

The advent of genetic tools to target specific cell-types has allowed scientists to ask meaningful questions related to brain cell physiology *in vivo*. Adeno-associated viruses (AAVs) allow discrete populations of cells in a specific brain region to be transfected with a gene of interest and subsequently, these cells can express the gene of interest (Kwon and Schafer, 2008; Burger et al., 2004). Coupled with cell-specific promoters which allow only a specific cell type to express the gene of interest, AAVs can be made for astrocytes and neurons to express a wide array of genetic products. Furthermore, with the addition of genetically coded sub-cellular tags, a particular protein of interest can be targeted to a discrete sub-cellular compartment. In astrocytes, it has been possible to target a green fluorescent protein to the plasma membrane using an Lck protein tyrosine kinase peptide sequence in order to visualize the entire bushy morphology of the cell (Benediktsson et al., 2005; Shigetomi et al., 2013).

With these technical advents in mind, a genetic toolkit that allows proteins to be captured and identified cell-specifically and within subcellular compartments would be ideal to dissect the protein underpinnings of astrocytes and neurons. Enzyme-mediated ligation methods have been used for decades to understand protein-protein interactions and their biological mechanisms. Most

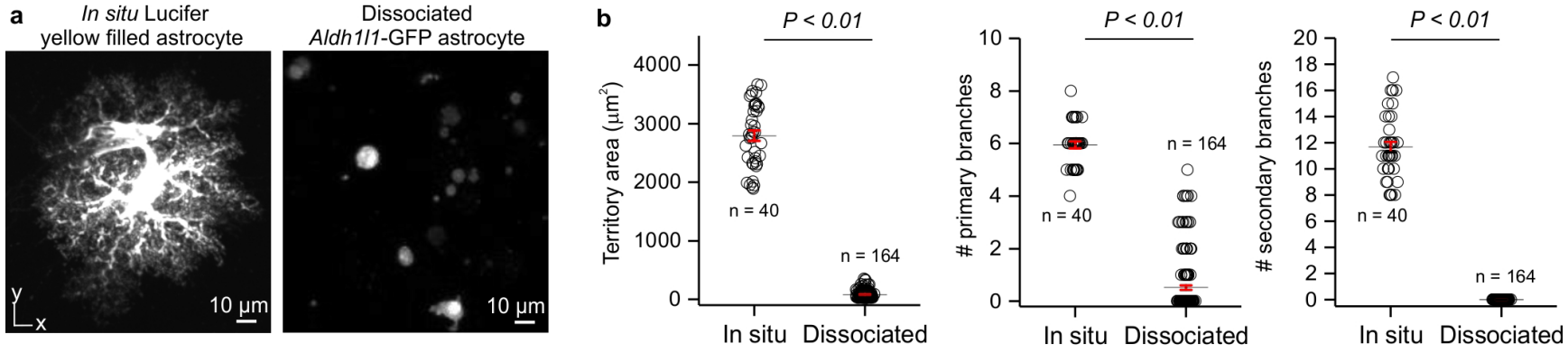

Figure 1.2: Cell dissociation leads to loss of astrocyte complex morphology. (a) *In situ* Lucifer yellow iontophoretically filled striatal astrocyte. Right panel shows GFP+ striatal astrocytes after a dissociation preparation for fluorescence activated cell sorting (FACS). **(b)** Quantification of territory area, number of primary branches, and number of secondary branches showed that GFP+ dissociated astrocytes have more decreased cellular complexity than Lucifer yellow dyed astrocytes (n = 40 cells from 15 mice for Lucifer yellow and n = 164 cells from 8 mice for *Aldh1l1*-GFP, respectively; territory area and number of primary branches: two-tailed student's unpaired t-test, P = 0.0002; number of secondary branches: two-tailed one-sample t-test, P = 0.003)

recently, the proteins that compose the inhibitory post-synapse in the hippocampus were characterized *in vivo* with a biotin-ligase called iBioID (Uezu et al., 2016). As a proof of concept, this biotin ligase was selectively expressed in the post-synaptic densities of neurons using AAV. Upon the addition of exogenous biotin, BioID labels proteins within a 10nm radius with biotin at free lysines. These biotinylated proteins can then be isolated with streptavidin, a strong biotin binder. These sampled proteins can then be identified with mass-spectrometry based proteomics (Roux et al., 2018; Sears et al., 2019). Novel BioID methods have been reported, including the development of BioID2.

A key feature for BioID to work in sub-cellular compartments is for its size to be as small and efficient as possible. This is because a larger enzyme is not trafficked along the cell as well as a smaller enzyme. While iBioID is based on proximity dependent biotinylation by a biotin ligase isolated from *Escherichia coli* and is 321 amino acids large, BioID2 is a smaller 233 amino acid long biotin ligase isolated from *Aquifex aeolicus* (Kim et al., 2016). Both its small size and higher biotinylation efficiency at lower concentrations of exogenous biotin are ideal for studying proteomes at distinct cellular compartments.

Section 1.6: Hypothesis and rationale

The central theme of this dissertation was to assess functional differences of astrocytes and neurons in a well described circuit, the striatum, at the proteomic level *in vivo*. This had not been previously explored and is the next frontier of glial biology. While we know that astrocytes and neurons are functionally distinct, a ground molecular basis of this distinction had not been assessed or addressed. Furthermore, functional differences between sub-cellular compartments had never been explored in either cell type. Using genetically-targeted BioID2 AAVs, we sought to obtain

and identify proteomes of astrocytes and neurons at the cytosol and at the plasma membrane (**Figure 1.3**). We then went further and assessed the sub-cellular proteomes of astrocytes by targeting BioID2 to five distinct genetically defined functional sub-compartments where astrocytes form functional contacts with other cells (**Figure 1.4**). These functional subcompartments are: (1) the astrocyte endfoot, which is the astrocyte interface with the brain vasculature using the Aquaporin-4 endfoot protein (Aqp4-BioID2), (2) the astrocyte gap junctions which are the interfaces at which astrocytes interact with other astrocytes using the Connexin-43 gap junction protein (Cx43-BioID2), (3) the astrocyte finest processes which interact with multiple cell types using actin regulatory protein, Ezrin (Ezr-BioID2), (4) astrocyte sites of glutamate buffering using astrocyte glutamate transporter protein Glt-1 (Glt-1-BioID2), and (5) astrocyte sites of potassium homeostasis using the astrocyte inward rectifying K^+ channel, Kir4.1 (Kir4.1-BioID2).

I hypothesized that key protein differences between astrocytes and neurons would be more prominent at the plasma membrane. Due to their distinct ionic properties, I hypothesized that astrocytes and neurons would express different ionic channels and receptors that reflect their differences in excitability (or lack thereof). Furthermore, proteins found within astrocytes sub-compartments could show major differences when compared to one another. Using the data we generated, I discovered that obsessive-compulsive disorder related protein, SAPAP3 (gene: *Dlgap3)*, is highly expressed in astrocytes and in select subcompartments including the astrocyte plasma membrane and the astrocyte fine processes where astrocytes abut synapses with Ezr-BioID2. I then utilized these proteomes to understand contributions of these cells to striatal physiology and relevant behavioral phenotypes, including behavior related to obsessive-compulsive disorder. Assessment of astrocyte and neuron proteomes in a circuit specific manner does not only allow us to understand fundamental functions and interactions of astrocytes and

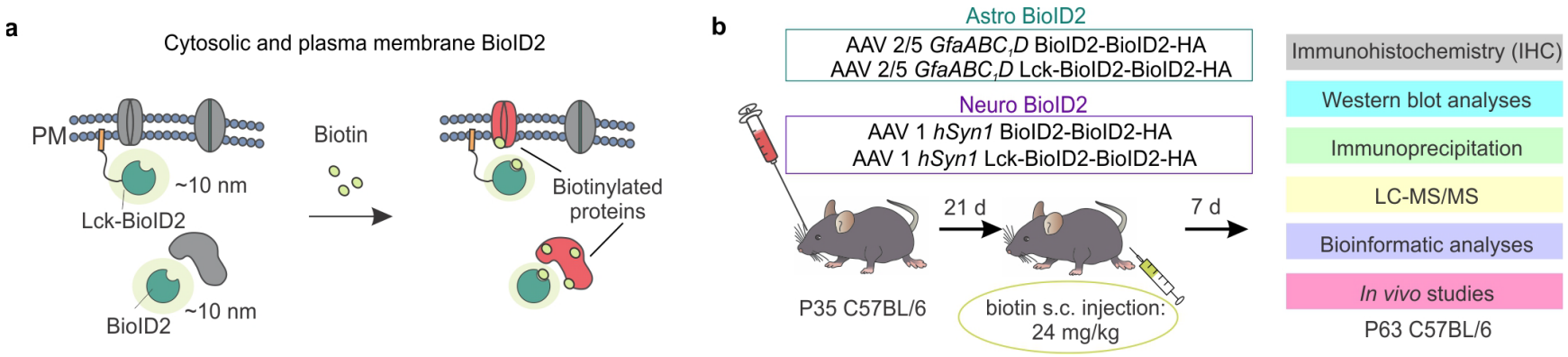

Figure 1.3: Workflow to assess cell- and subcompartment specific proteomes *in vivo*. (a) Diagram of BioID2 biotin labeling in mammalian cells. Proximal (~10nm) proteins are biotinylated upon the addition of exogenous biotin. **(b)** Schematic of the cell and compartment specific *in vivo* BioID2 experimental design using adeno-associated viruses (AAVs).

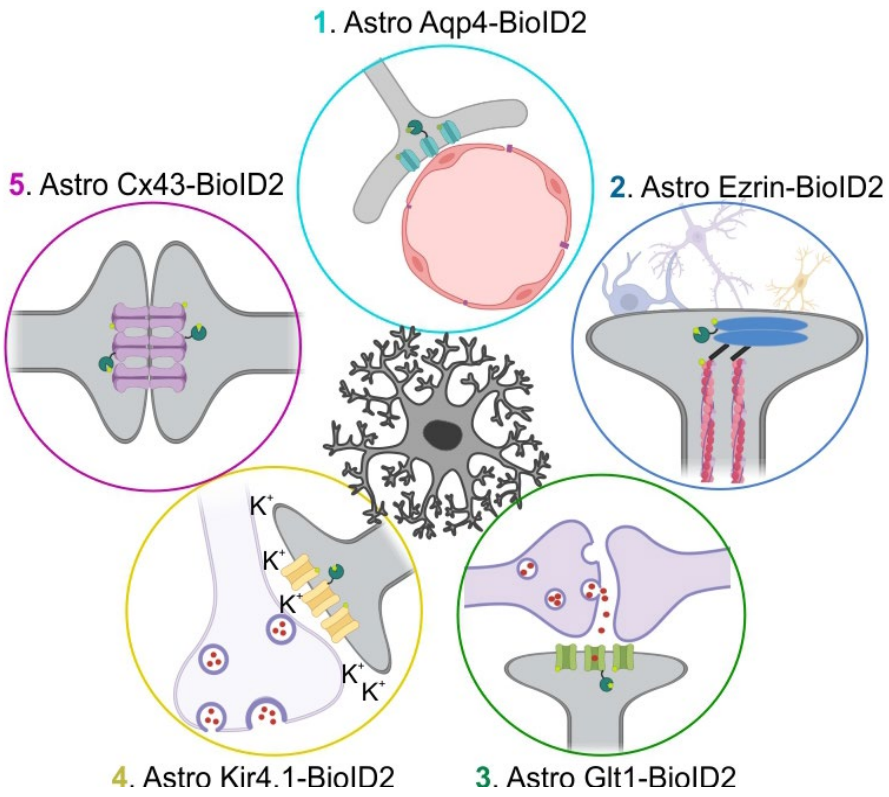

Figure 1.4: Astrocyte subcompartment specific BioID2. Cartoon illustrating five astrocyte subcompartments that were genetically targeted with BioID2.

neurons, but also allows a discovery platform for treatments of specific psychiatric or degenerative diseases.

Section 1.7: Summary

Ever since Virchow's famed lecture to a group of medical students a century ago, the implication that astrocytes are dynamic contributors to brain function and dysfunction has been largely explored, but discrete mechanisms remain to be elucidated. While much progress has been made on astrocyte-neuron interactions in the striatum, assessing the protein profiles that allow these interactions to occur had not been done. The goal of identifying cell- and subcompartment-specific proteomes of neurons and astrocytes in the striatum was to identify the molecularly functional correlates that give rise to cell function. Furthermore, exploration of these proteomes gave rise to a range of diverse hypotheses, one of which we tested in the context of obsessive compulsive disorder.

While we focused on the striatum, the palette of targeted BioID2 constructs for astrocytes and neurons built in this dissertation can be of use for other brain regions and in a variety of different contexts. In this way, a systematic analysis of functional proteomes allows us to go one step further in our understanding of astrocytes in complex neural circuits.

CHAPTER 2: METHODS

Section 2.1: Experimental model and subject details

Cell lines

HEK293 cells (sex: female, RRID: CVCL_0045) were obtained from ATCC and maintained in 25cm^2 cell culture flasks in DMEM/F12 media with Glutamax (Invitrogen) supplemented with 10% fetal bovine serum and penicillin/streptomycin. Cells were grown in a humidified cell culture incubation with 95% air/5% CO_2 at 37°C.

Mouse models

All experiments were conducted in accordance with the National Institutes of Health (NIH) Guide for the Care and Use of Laboratory Animals and were approved by the Chancellor's Animal Research Committee at the University of California, Los Angeles (UCLA). Male and female mice aged between 3 weeks and 28 weeks were used in this study. Mice were housed in the vivarium managed by the UCLA Division of Laboratory Animal Medicine with a 12h light/dark cycle and *ad libitum* access to food and water. Wild-type C57BL/6NTac mice were purchased from Taconic Biosciences. Targeted knockout mice for *DLGAP3*/SAPAP3 (B6.129-Dlgap3tm1Gfng/J) were obtained from Jackson laboratories and maintained as a heterozygous line (Welch et al., 2007).

All animal experiments throughout this dissertation represent a minimum of three animals per group replicated at least two times. For injection of adeno-associated viruses (AAVs), mice were randomly assigned to each AAV construct group. Alternating batches of female and male mice were used when possible and when available.

Section 2.2: Cloning and generation of adeno-associated viruses (AAVs)

The BioID2 (BirA) sequence with an HA tag was PCR amplified from pAAV-BioID2-Linker-BioID2-HA plasmid (a kind gift from Dr. Scott Soderling) and incorporated into pZac2.1 vector using the In-fusion cloning kit (Takara Bio) to generate plasmids, PZac2.1 GfaABC$_1$D BioID2-Linker-BioID2-HA (Astro BioID2) and PZac2.1 hSyn1 BioID2-Linker-BioID2-HA (Neuro BioID2) which carry the astrocyte-specific promoter and the human Synapsin-1 neuronal promoter, respectively.

For generation of the sub-compartment specific BioID2 constructs, the Astro BioID2 plasmid was cut by restriction digest at the *XhoI* site. Sequences for each sub-compartment were amplified by PCR, and the sequences were cloned into the *XhoI* site using the In-fusion cloning kit (Takara Bio). The Aqp-4 sequence was PCR amplified from GeneCopoeia plasmid EX-Mm20326; the Cx-43 sequence was amplified from msfGFP-Cx43 plasmid (Addgene #69024); the GLT-1 sequence was amplified from GeneCopoeia plasmid Mm27106; the Ezr sequence was amplified from GeneCopoeia plasmid Mm2129m; the Kir4.1 sequence was amplified from GfaABC$_1$D-eGFP-Kir4.1 plasmid (Addgene #52874). The sub-compartment GFP controls were generated by digesting the resulting sub-compartment BioID2 plasmids with *BamHI,* which removed the BioID2 sequence, and inserting PCR amplified eGFP sequence from GfaABC$_1$D-eGFP-Kir4.1 plasmid (Addgene #52874) using Infusion cloning kit (Takara Bio).

To generate the eGFP-SAPAP3 and the HA-BioID2-SAPAP3 AAVs, the SAPAP3 sequence was amplified from GeneCopoeia plasmid Mm16264. PZac2.1-gfaABC$_1$D-BioID2-Linker-BioID2-HA and PZac2.1 hSyn1 BioID2-Linker-BioID2-HA were digested with *BamHI* and the SAPAP3 sequence was cloned into the *BamHI* sites using In-fusion cloning kit (Takara Bio) which generated PZac2.1 GfaABC$_1$D SAPAP3 and PZac2.1 hSyn1 SAPAP3. This resulting plasmid was

then digested with *BmtI*. eGFP sequence was PCR amplified from GfaABC$_1$D-eGFP-Kir4.1 plasmid (Addgene #52874) and cloned into the PZac2.1 SAPAP3 plasmids at the *BmtI* site with In-fusion cloning kit (Takara Bio) producing PZac2.1 GfaABC$_1$D eGFP-SAPAP3 and PZac2.1 hSyn1 eGFP-SAPAP3. The HA-BioID2 sequence was PCR amplified from an existing PZac2.1 plasmid and cloned into the PZac2.1 SAPAP3 plasmids at the *BmtI* site using the Infusion cloning kit (Takara Bio) producing PZac2.1 GfaABC$_1$D HA-BioID2-SAPAP3 and PZac2.1 Syn1 HA-BioID2-SAPAP3. Generation of pZac2.1 *GfaABC$_1$D* 3xHA-SAPAP3 consisted of digesting the pZac2.1 *GfaABC$_1$D* SAPAP3 plasmid with *BmtI* and ligating an annealed oligo with the 3xHA sequence into the *BmtI* sites using T4 ligase.

To generate the astrocyte specific LifeAct-eGFP AAV, the LifeAct-eGFP sequence was amplified from pEGFP-C1 LifeAct-EGFP (Addgene #58470). pZac2.1 *GfaABC$_1$D* BioID2-Linker-BioID2-HA plasmid was digested with *BamHI*, and the LifeAct-eGFP sequence was cloned into the *BamHI* sites using In-fusion cloning kit (Takara Bio) producing pZac2.1 *GfaABC$_1$D* LifeAct-eGFP.

All plasmids generated in this study are listed in **Table 2.1**. All BioID2 plasmids and their GFP counterparts were sent to the Penn Vector core for AAV production. All SAPAP3 associated plasmids were sent to Virovek for AAV production. Astrocyte specific constructs were produced in AAV2/5 serotype while neuron specific constructs were produced in AAV1 serotype. AAV titers can be found in **Table 2.1**.

For HEK293 cell transfection, the BioID2 sequence was cloned into pcDNA3.1 vector between *BamHI* sites to generate pcDNA3.1-CMV-BioID2-Linker-BioID2-HA and pcDNA3.1-CMV-Lck-BioID2-Linker-BioID2-HA.

AAV plasmid construct	Addgene ID	AAV Titer
*AAV2/5 *GfaABC$_1$D* BioID2-Linker-BioID2-HA	176740	3.7 x 10^{13} gc/ml
*AAV2/5 *GfaABC$_1$D* Lck-BioID2-Linker-BioID2-HA	176741	3.1 x 10^{13} gc/ml
*AAV2/5 *GfaABC$_1$D* Aqp4-BioID2-Linker-BioID2-HA	176742	1.1 x 10^{13} gc/ml
*AAV2/5 *GfaABC$_1$D* Cx43-BioID2-Linker-BioID2-HA	176743	2.5 x 10^{13} gc/ml
*AAV2/5 *GfaABC$_1$D* Ezr-BioID2-Linker-BioID2-HA	176744	1.0 x 10^{13} gc/ml
*AAV2/5 *GfaABC$_1$D* Glt1-BioID2-Linker-BioID2-HA	176745	2.0 x 10^{13} gc/ml
*AAV2/5 *GfaABC$_1$D* HA-BioID2-Linker-BioID2-Kir4.1	176746	6.3 x 10^{13} gc/ml
AAV2/5 *GfaABC$_1$D* tdTomato	44332	1.4 x 10^{13} gc/ml
AAV2/5 *GfaABC$_1$D* Lck-eGFP	52925	1.5 x 10^{13} gc/ml
*AAV2/5 *GfaABC$_1$D* eGFP	176861	1.1 x 10^{13} gc/ml
*AAV2/5 *GfaABC$_1$D* Aqp4-eGFP	176747	2.4 x 10^{13} gc/ml
*AAV2/5 *GfaABC$_1$D* Cx43-eGFP	176860	1.4 x 10^{13} gc/ml
*AAV2/5 *GfaABC$_1$D* Ezr-eGFP	176856	1.6 x 10^{13} gc/ml
*AAV2/5 *GfaABC$_1$D* Glt1-eGFP	176862	5.8 x 10^{13} gc/ml
AAV2/5 *GfaABC$_1$D* eGFP-Kir4.1	52874	2.7 x 10^{12} gc/ml
*AAV5 *GfaABC$_1$D* eGFP-SAPAP3	176859	1.0 x 10^{13} gc/ml
*AAV5 *GfaABC$_1$D* HA-BioID2-SAPAP3	176858	2.4 x 10^{13} gc/ml
AAV2/5 *GfaABC$_1$D* Rpl22-HA	111811	2.1 x 10^{13} gc/ml
*AAV5 *GfaABC$_1$D* HA-SAPAP3	190200	1.0 x 10^{13} gc/ml
*AAV5 *GfaABC$_1$D* LifeAct-eGFP	190199	1.0 x 10^{13} gc/ml
*AAV1 *hSyn1* HA-BioID2-SAPAP3	176857	2.0 x 10^{13} gc/ml
*AAV1 *hSyn1* eGFP-SAPAP3	176853	2.0 x 10^{13} gc/ml
*AAV5 *hSyn1* Rpl22-HA	177685	1.0 x 10^{13} gc/ml
*AAV1 *hSyn1* eGFP	50465	1.3 x 10^{13} gc/ml

Table 2.1. Plasmids for AAV production used in this study (* new AAV plasmids generated in this dissertation).

Section 2.3: HEK293 cell experiments

HEK293 cell immunocytochemistry

HEK293 cells were prepared for transfection by plating onto six-well plates and transfection was performed when cells reached 80% confluence. For expression in HEK293 cells, 0.4 µg plasmid DNA was transfected using Effectene transfection reagent (QIAGEN). Biotin (Sigma B4501) was dissolved in sterile 0.1M PBS to make a 1mM stock solution. The stock solution was added to the HEK293 cell media for a final concentration of 50 mM biotin. 48 hours post-transfection, the cells were trypsinized and were transferred to poly-D-lysine coverslips. 24 hours later the cells were washed once with 0.1 M phosphate buffered saline (PBS) and then fixed with 10% formalin for 10 minutes. Cells were washed in 0.1M PBS and then incubated with agitation in a blocking solution containing 5% NGS in 0.1M PBS with 0.2% Triton-X (Sigma) in PBS for 1 hr at room temperature. The cells were then incubated with agitation in mouse anti-HA primary antibody (1:1,000; Biolegend, 901514) diluted in 0.1M PBS with 5% NGS at 4°C overnight. Cells were then incubated with agitation with the following secondary antibodies and fluorophores in a solution containing 5% NGS in 0.1M PBS for 2 hrs at room temperature (1:1,000; Molecular probes): Alexa Fluor 546 goat anti-mouse (A11003), Streptavidin, Alexa Fluor 488 conjugate (S11223). The cells were rinsed 3 times in 0.1M PBS for 10 minutes each. The coverslips containing the cells were then mounted onto microscope slides in fluoromount-G. Fluorescent images were taken using UplanSApo 40x 1.30 NA oil immersion objective lens on a confocal laser-scanning microscope (FV10-ASW; Olympus). Laser settings were the same for all cells. Images represent maximum intensity projections of optical sections with a step size of 1.0 µm.

HEK293 cell western blots

For western blotting the cells were directly lysed in RIPA buffer (150 mM NaCl, 50 mM Tris pH 8.0, 1% Triton-X, 0.5% Na Deoxycholate, 0.1% SDS, and Halt protease inhibitor (Thermo Scientific 78429). The cells were homogenized using a cell scratcher and the lysate was incubated at 4°C while rotating for 30 mins. The samples were sonicated for 10 minutes and then centrifuged at 16500 x g for 10 minutes at 4°C. The supernatant was collected and the protein concentrations were measured using the BCA protein assay (Thermo scientific). The samples were then mixed with 2x laemmli solution (BioRad) containing β-mercaptoethanol. The samples were boiled at 95°C for 10 minutes before being electrophoretically separated by 10% SDS-PAGE (30 µg protein per lane) and transferred unto nitrocellulose membrane (0.45 µm). The membrane was incubated with agitation in a solution containing 5% BSA, 0.1% Tween-20, and 0.1M PBS for 1 hr. The membrane was probed with streptavidin-HRP (Sigma RABHRP3) at 1:250 for two hours. The membrane was then treated with the Pierce chemiluminescence solution for 1 min and imaged. The blot was incubated overnight at 4°C with rabbit anti-β-actin (1:1,1000; abcam ab8227). IR-dye 800CW anti-rabbit (1:10,000; Li-Cor) secondary was used and images were acquired on a Li-Cor odyssey infrared imager. Signal intensities at expected molecular weight were quantified using ImageJ. The streptavidin signal levels were normalized to β-actin by dividing the streptavidin signal intensity by the β-actin signal intensity.

Section 2.4: Stereotaxic microinjections of adeno-associated viruses (AAVs)

Microinjections for proteomic and RNA-sequencing studies

All surgical procedures were conducted under general anesthesia using continuous isoflurane (induction at 5%, maintenance at 2% vol/vol) in 6 week old C57/BL6NTac mice unless otherwise

stated. Anesthetic depth was continuously monitored and adjusted when necessary. After induction of anesthesia, mice were fitted into the stereotaxic frame (David Kopf Instruments), their noses placed into a veterinary grade anesthesia ventilation system (VetEquip) and their heads were secured by blunt ear bars. Mice were administered 0.1 mg/kg of buprenorphine (Bupranex) subcutaneously before surgery. The surgical incision site was cleaned three times with 10% povidone iodine and 70% ethanol (vol/vol). A skin incision was made followed by craniotomies (1-2 mm in diameter) above the left parietal cortex using a small steel burr (NeoBurr) powered by a high speed drill (Midwest Tradition). Sterile saline (0.9%) was applied onto the skull to reduce heating caused by drilling. One craniotomy was made for unilateral injections, while two craniotomies were made for bilateral injections. The injections were carried out by using the stereotaxic apparatus to guide the placement of beveled glass pipettes (1B100-4, World Precision Instruments). For the left striatum the coordinates were: 0.8mm anterior to bregma, 2mm lateral to the midline, and 2.4 mm from the pial surface. Adeno-associated virus (AAV) was injected by using a syringe pump (SmartTouch Pump, World Precision Instruments).

Following AAV microinjection, the glass pipette was left in place for at least 10 min prior to slow withdrawal. Surgical wounds were closed with external 5-0 nylon sutures. Following surgery, animals were allowed to recover overnight in cages placed on a low-voltage heating pad. Buprenorphine was administered two times a day for 48 hrs after surgery. Thimethoprim sulfamethoxazole was provided in food to the mice for one week. Virus injected mice were used for experiments at least three weeks post-surgery.

All AAV titers were adjusted to 1.0 x 10^{13} genome copies/mL with sterile 0.1 M phosphate buffered saline (PBS). The viruses used were: 0.5 µl AAV2/5 GfaABC$_1$D BioID2-Linker-BioID2-HA, 0.5 µl AAV2/5 GfaABC$_1$D Lck-BioID2-Linker-BioID2-HA, 0.5 µl AAV2/5 GfaABC$_1$D

Aqp4-BioID2-Linker-BioID2-HA, 0.5 µl AAV2/5 GfaABC$_1$D Cx43-BioID2-Linker-BioID2-HA, 0.5 µl AAV2/5 GfaABC$_1$D Ezr-BioID2-Linker-BioID2-HA, 0.5 µl AAV2/5 GfaABC$_1$D Glt1-BioID2-Linker-BioID2-HA, 0.5 µl AAV2/5 GfaABC$_1$D HA-BioID2-Linker-BioID2-Kir4.1, 0.5 µl AAV2/5 GfaABC$_1$D tdT (Addgene #44332-AAV5), 0.5 µl AAV2/5 GfaABC$_1$D Lck-eGFP (Addgene #105598-AAV5), 0.5 µl AAV2/5 GfaABC$_1$D Aqp4-eGFP, 0.5 µl AAV2/5 GfaABC$_1$D Cx43-eGFP, 0.5 µl AAV2/5 GfaABC$_1$D Ezr-eGFP, 0.5 µl AAV2/5 GfaABC$_1$D Glt1-eGFP, 0.5 µl AAV2/5 GfaABC$_1$D eGFP-Kir4.1, 0.5 µl AAV1 hSyn1 BioID2-Linker-BioID2-HA, 0.5 µl AAV1 hSyn1 Lck-BioID2-Linker-BioID2-HA, 0.5 µl AAV1 hSyn1 eGFP (Addgene #50465-AAV1), 0.5 uL AAV1 hSyn1 Lck-GFP, 0.5 µl AAV2/5 GfaABC$_1$D Rpl22-HA (Addgene #111811), and 0.5 µl AAV1 hSyn1 Rpl22-HA. For co-immunoprecipitation experiments, 0.1 µl of *GfaABC$_1$D* Ezr-eGFP was co-injected with 0.1 µl of *GfaABC$_1$D* 3xHA-SAPAP3 (Addgene #190200) and 0.1 µl of *GfaABC$_1$D* Glt1-eGFP was co-injected with 0.1 µl of *GfaABC$_1$D* 3xHA-SAPAP3.

Microinjections for studies of SAPAP3

Surgical procedures for SAPAP3 KO mice were conducted as described above. Briefly, GFP, GFP-SAPAP3 AAV, LifeAct-GFP, or tdTomato was bilaterally injected into the striatum of 3-4 week old mice through 2 sites at 3 locations per hemisphere. At each of the injection sites, the microinjection needle was advanced to the deepest (ventral) position for the first injection while the additional injections were made every 0.3 mm while withdrawing the injection needle. The coordinates from bregma were: injection site 1: anterior 0.5 mm, mediolateral 1.5mm, dorsoventral 2.9 mm, 2.6 mm, 2.3 mm from pial surface; injection site 2: anterior 0.5 mm, mediolateral -1.5mm, dorsoventral 2.9 mm, 2.6 mm, 2.3 mm from pial surface.

For each injection location, 150 nl of virus was injected and the needle was left in place for 5 min after each injection. All AAV titers were adjusted to 1.0 x 1013 genome copies/ml with sterile 0.1 M phosphate buffered saline (PBS). The viruses used were: AAV 2/5 GfaABC1D LifeAct-eGFP (Addgene # 190199), AAV2/5 GfaABC$_1$D tdT (Addgene #44332-AAV5), AAV2/5 GfaABC1D eGFP-SAPAP3, AAV2/5 GfaABC1D eGFP, AAV1 hSyn1 eGFP-SAPAP3, AAV1 hSyn1 eGFP (Addgene #50465-AAV1). Litters with multiple SAPAP3 KO mice were split between experimental groups.

Section 2.5: *in vivo* BioID2 protein biotinylation

3 weeks post-AAV microinjection, mice were treated with a subcutaneous injection of biotin at 24mg/kg (Millipore Sigma RES1052B-B7) dissolved in sterile 0.1 M phosphate buffered saline (PBS) once per day for 7 consecutive days. The mice were processed 16 hours after the last biotin injection.

Section 2.6: Immunohistochemistry of floating sections

Mice were transcardially perfused with chilled 0.1 M phosphate buffered saline (PBS) followed by 10% formalin. After gentle removal of the skull, the brains were post-fixed in 10% formalin for 6 hours. The brains were then cryoprotected in 30% sucrose with 0.1 M PBS solution for at least 48 hours at 4°C. Serial coronal sections (40 μm) containing striatum were prepared using a cryostat microtome (Leica) at -20°C and processed for immunohistochemistry. Sections were washed 3 times in 0.1 M PBS for 10 min each and then incubated in a blocking solution consisting of 5% NGS in 0.1 M PBS with 0.2% Triton-X for 1 hr at room temperature with agitation. Sections were then incubated in primary antibodies diluted in 5% NGS in 0.1 M PBS solution overnight at 4°C.

The following primary antibodies were used: mouse anti-HA (1:1,000; Biolegend, 901514), rabbit anti-HA (1:1000; abcam ab9110), rabbit anti-S100β (1:1,000; Abcam ab13970), rabbit anti-NeuN (1:1,000; Cell Signaling 12943S), guinea pig anti-neuN (1:1000; Synaptic Systems 266004), rabbit anti-DARPP32 (1:1,000; Abcam ab40801), guinea pig anti-DARPP32 (1:1000, Frontier Institute DARPP-Gp-A250), chicken anti-GFP (1:1,000; Abcam ab13970), rabbit anti-SAPAP3 (1:100; a kind gift from Dr. Guoping Feng), mouse anti-RFP (1:500; Rockland 600906379), rabbit anti-PAICS (1:100; Invitrogen 92985), mouse anti-Nebl (1:100; Santa Cruz Biotechnology 393784), rabbit anti-Slc4a4/NBC1 (1:100; Novus NBP32020), rabbit anti-Arpc1a (1:100; Invitrogen 102339), rabbit anti-Faim2 (1:100; Origen TP300196), rabbit anti-Hepacam (1:100; Novus Biologicals 04983), mouse anti- APC (1:500; abcam ab16794), rabbit anti- Olig2 (1:500; Millipore AB9610), rabbit anti- ΔFosB (1:500; Cell Signaling Technology 14695S).

Sections were then incubated with the following secondary antibodies for 2 hrs at room temperature (1:1,000; Molecular probes): Alexa Fluor 488 goat anti-chicken (A11039), Alexa Fluor 647 goat anti-rabbit (A21244), Alexa Fluor 546 goat anti-mouse (A11003), Alexa Fluor 488 goat anti-rabbit (A11008), Streptavidin, Alexa Fluor 488 conjugate (S11223).

The free floating sections were mounted on microscope slides in fluoromount-G. Fluorescent images were taken using UplanFL 40x 1.30 NA oil immersion or PlanApo N 60x 1.45 NA oil immersion objective lens on a confocal laser-scanning microscope (FV3000; Olympus). Laser setting were kept the same within each experiment. Images represent maximum intensity projections of optical sections with a step-size of 1.0 μm. Images were processed on ImageJ. Cell counting was done on maximum intensity projections using the Cell Counter plugin; only cells with somata completely in the region of interest were counted. Co-localization analysis was conducted with the ImageJ Coloc2 plugin.

Section 2.7: Dual *in situ* hybridization with RNAscope and IHC

Fixed-frozen tissue was processed as described above. Serial coronal sections (14 μm) containing striatum were prepared using a cryostat microtome (Leica) at -20°C and mounted immediately unto glass slides. Dual ISH-IHC was performed using the Multiplex RNAscope v2 with integrated co-detection work flow (ACDBio 323180 and 323110). Sections were baked for 30 min at 60°C. Sections were washed for at least 15 mins in 0.1 M phosphate buffered saline and then incubated in 1x Target Retrieval Reagents for 5 min at 95°C. After washing with ddH_2O twice, the sections were dehydrated with 100% ethanol and dried at room temperature. Sections were then incubated with primary antibodies rabbit anti- S100β (1:500; Abcam ab13970), rabbit anti-RFP to amplify astrocyte specific tdTomato signal (1:500; Rockland 600-401-379), guinea pig anti-NeuN (1:500; Synaptic Systems 266004) overnight at 4°C. Sections were then incubated with Protease Pretreat-4 solution (ACDBio 322340) for 30 min at 40°C. The sections were washed with ddH_2O twice for 1 min each and then incubated with probe for 2 hours at 40°C: Mm-DLGAP3-C3 (ACDBio 586091-C3), Mm-Mapt-C1 (ACDBio 400351), Mm-Tjp1-C1 (ACDBio 440411). The sections were incubated in Amp 1-FL for 30 min, AMP 2-FL for 15 min, AMP 3-FL for 30 min, and AMP 4-FL for 15 min at 40°C while washing in 1x wash buffer (ACDBio 310091) between incubations. The HRP-C3 signal was developed with Opal 690 fluorophore (Akoya Biosciences FP1497001KT). The HRP-C1 signal was developed with Opal 520 fluorophore (Akoya Biosciences FP1487001KT). All incubations at 40°C or 60°C were performed in the HybEZ hybridization System (ACDBio).

Lastly, sections were incubated with Alexa Fluor goat secondary antibodies described in the IHC section for 45 min at RT. Images were obtained in the same way as IHC as described above with a step size of 0.5 μm. Images were processed with ImageJ. Astrocyte somata were labeled

with S100β signal and number of puncta within each soma was measured. Astrocyte territories were labeled with tdTomato signal and the number of puncta within each territory was measured.

Section 2.8: Proximity ligase assay (PLA) for protein-protein interaction validation

Fixed frozen tissue was processed as described above. Serial coronal sections (20 μm) containing striatum sparsely labeled with astrocyte-specific tdTomato were prepared using a cryostat microtome (Leica) at -20°C and mounted immediately unto glass slides. The proximity ligation assay was performed using the Sigma-Aldrich Duolink PLA fluorescence protocol (Sigma-Aldrich DUO92101 and DUO92013). Sections were baked for 30 min at 60°C. Sections were washed for at least 15 mins in 0.1 M phosphate buffered saline (PBS). After washing, sections were incubated in 1x citrate pH 6.0 antigen retrieval buffer (Sigma C999) for 10 min at 90°C. After washing three times in 0.2% Triton-X in PBS (PBS-T), the sections were blocked for 45 min at RT with 5% Donkey serum (Sigma D9663) in PBS-T. Sections were then incubated with primary antibodies overnight at 4°C: rabbit anti- SAPAP3 (1:50), mouse anti- Glt1 (1:50; Santa Cruz Biotechnology sc-365634), mouse anti- Ezrin (1:100; BioLegend 866401), guinea pig anti- RFP (1:500; Synaptic Systems 390004). Sections were then incubated with PLA probe cocktail containing the anti-rabbit PLUS primer probe (DUO92002) and the anti-mouse MINUS primer probe (DUO92004) for 1 hr at 37°C. The sections were washed twice in 1X wash buffer A (DUO82049). Sections were then incubated with ligation solution containing ligase for 30 min at 37°C. Sections were once again washed twice with 1X wash buffer A and then incubated with amplification solution containing DNA polymerase for at least 3 hours at 37°C. Sections were then washed twice in 1X wash buffer B (DUO82049) and then washed in 0.01X wash buffer B. To amplify the tdTomato signal, sections were then incubated with donkey anti-guinea pig Cy3

(1:500; Jackson ImmunoResearch 706-165-148) for 45 min at RT. Sections were washed twice with PBS and then coverslips were mounted with DuoLink mounting medium with DAPI (DUO82040). Images were obtained in the same way as IHC as described above with a step size of 0.5 μm. Images were processed with ImageJ. Astrocyte territories were labeled with tdTomato signal and number of puncta and integrated intensity within each territory was measured.

Section 2.9: Western blots of *in vivo* BioID2 samples

Mice were decapitated and striata were dissected and homogenized with a dounce and pestle in ice cold RIPA buffer (150 mM NaCl, 50 mM Tris pH 8.0, 1% Triton-X, 0.5% Na Deoxycholate, 0.1% SDS, and Halt protease inhibitor (Thermo Scientific 78429). The homogenate was incubated at 4°C while spinning for 1 hr. The homogenate was sonicated and then centrifuged at 4°C for 10 minutes at 15600 xg. The clarified lysate was collected and the protein concentration was measured using the BCA protein assay (Thermo). The samples were then processed as described for HEK cells above and analyzed as stated above with 30 ug of protein loaded into each gel well.

Section 2.10: Co-immunoprecipitation of SAPAP3 interacting proteins

Mice injected with either combination in the striatum: *GfaABC$_1$D* 3xHA-SAPAP3 and *GfaABC$_1$D* Glt1-eGFP, *GfaABC$_1$D* 3xHA-SAPAP3 and *GfaABC$_1$D* Ezr-eGFP, *GfaABC$_1$D* 3xHA-SAPAP3 only, *GfaABC$_1$D* Glt1-eGFP only, or *GfaABC$_1$D* Ezr-eGFP only were decapitated and striata were dissected and homogenized with a dounce and pestle in ice cold lysis buffer (25 mM HEPES pH 7.5, 150 mM NaCl, 1 mM EDTA, 1% NP-40, 5 mM NaF, 1 mM orthovanadate, and Halt protease inhibitor cocktail (Thermo Scientific 78429)). The homogenate was incubated at 4°C while spinning for 1 hr. The homogenate was then centrifuged at 4°C for 15 min at 15,000 x g.

The supernatant was further cleared by ultracentrifugation at 100,000 x g for 30 min at 4°C. The cleared lysate was then incubated with GFP-trap beads (chromotek #gtma) or incubated with anti-HA tag beads (Thermo #88836) overnight at 4°C. The beads were then washed 3 times with wash buffer (25mM HEPES pH 7.5, 500 mM NaCl, 1mM EDTA, 1% NP-40, 5 mM NaF, 1 mM orthovanadate, and Halt protease inhibitor cocktail). 1x Laemmli buffer was prepared with β-mercaptoethanol (BioRad #1610737) and added to the beads. The beads were boiled in the Laemmli buffer for 10 min at 95°C. The bead supernatants were cooled and loaded on an SDS-PAGE gel for western blot analyses as described above. Primary antibodies used were: chicken anti-GFP (1:1,000; Abcam ab13970), rabbit anti-HA (1:1000; abcam ab9110). Secondary antibodies used were: goat anti-rabbit plus 647 (1: 2000; Invitrogen A32733) and goat anti-chicken plus 555 (1: 2000; Invitrogen A32932).

Section 2.11: Whole tissue protein extraction for mass spectrometry

Striata from 6 month old SAPAP3 WT or KO mice were lysed in 200 µl lysis buffer containing 8 M urea, 50 mM Tris-HCl pH 8.2, 75 mM NaCl, 5 mM EDTA, 5 mM EGTA, 10 mM sodium pyrophosphate, and protease inhibitor cocktail. Tissue was dounce homogenized sonicated for 10 minutes at 80% power in a bath soniciator. Samples were centrifuged at 15,000 x g for 20 min at 4°C to remove debris. The supernatant was collected and then processed for downstream mass spectrometry analysis.

Section 2.12: *in vivo* BioID2 biotinylated protein pulldown

Purification of biotinylated proteins was conducted as previously described (Uezu et al., 2016). Each AAV BioID2 probe and its counterpart AAV GFP control were injected into the striatum of

6 week old C57/BL6NTac mice. 3 weeks post-AAV microinjection, biotin (Millipore Sigma RES1052B-B7) was subcutaneously injected at 24 mg/kg for 7 consecutive days. All mice were processed 16 hours after the last biotin injection. 8 mice were used for each biotinylated protein purification and each purification was performed independently 4 times. Mice were decapitated and striata were microdissected. Striata from 4 mice were dounce homogenized with 600 µl of lysis buffer 1 (1 mM EDTA, 150 mM NaCl, 50 mM HEPES pH 7.5 supplemented with Halt protease inhibitor (Thermo Scientific 78429). Immediately after homogenization, 600 µl of lysis buffer 2 (2% Na Deoxycholate, 2% Triton-X, 0.5% SDS, 1mM EDTA, 150 mM NaCl and 50 mM HEPES pH 7.5) was added. The lysed samples were sonicated for 5 min at 60% power and then centrifuged at 15,000 x g for 15 min at 4°C. The resulting supernatant was then ultracentrifuged at 100,000 x g for 30 min at 4°C. SDS was added to the supernatant for a final concentration of 1%. The sample was then boiled at 95°C for 5 minutes. Sample was cooled on ice and incubated with 35 µl of equilibrated anti-pyruvate carboxylase (5 ug; abcam 110314) conjugated agarose beads (Pierce 20398) for 4 hrs at 4°C while rotating. Subsequently, the sample was centrifuged at 2000 rpm for 5 mins at 4°C and the supernatant was incubated with 80 µl NeutrAvidin agarose at 4°C overnight while rotating.

The NeutrAvidin beads were then washed twice with 0.2% SDS, twice with wash buffer (1% Na deoxycholate, 1% Triton-X, 25 mM LiCl), twice with 1M NaCl, and 5 times with 50 mM ammonium bicarbonate. Proteins bound to the agarose were then eluted in elution buffer (5 mM biotin, 0.1% Rapigest SF surfactant, 50 mM ammonium bicarbonate) at 60°C for a minimum of 2 hr.

The final protein concentration was measured by BCA. Eluted proteins were subject to reduction using 5mM Tris (2-carboxyethyl) phosphine for 30 min and then alkylated by 10 mM

iodoacetamide for another 30 min. Proteins were digested with Lys-C for 4 hrs at a 1:100 protease-to-peptide ratio overnight. The digestion reaction was terminated with the addition of formic acid to 5% (vol./vol.) with centrifugation. Each sample was then desalted with C18 tips (Thermo Scientific, 87784) and dried with a SpeedVac vacuum concentrator. The peptide pellet was reconstituted in 5% formic acid before liquid chromatography-tandem mass spectrometry (LC-MS/MS). A schematic of the biotinylated protein pulldown procedure is provided as **Figure 2.1**.

Section 2.13: Quantitative mass spectrometry analysis of biotinylated proteins

Tryptic peptide mixtures were fractionated online using a 25 cm long, 75μm ID fused-silica capillary that was packed in-house with bulk ReproSil-Pur 120 C18-AQ particles as described previously (Jami-Alahmadi et al., 2021). Peptides were analyzed using a 140-min water-acetonitrile gradient. The solvent gradient began with LC buffer B (acetonitrile solution with 3% DMSO and 0.1% formic acid) to LC buffer A (water solution with 3% DMSO and 0.1% formic acid) increasing from 0 to 6% over 5 min at a flow rate of 400 nL/min. The flow rate was then reduced to 200 nL/min and a linear gradient beginning at 6-28% LC buffer B was applied to the column over the course of 123 min. The linear gradient of LC buffer B was then further increased to 28-35% for 8 min followed by a rapid ramp-up to 85% for column washing using a Dionex Ultimate 3000 UHPLC (Thermo Fisher Scientific). The eluted peptides were then ionized via nanoelectrospray ionization and entered the Orbitrap Fusion Lumos high resolution accurate tandem mass spectrometer (Thermo Fisher Scientific) with an MS1 resolution (*r*) of 120,000 followed by sequential MS2 scans at a resolution (*r*) of 15,000. Data generated by LC-MS/MS was processed in two discrete bioinformatic pipelines. The first analysis used the Integrated Proteomics Pipeline 2 (Integrated Proteomics Applications; San Diego, CA) to generate spectral

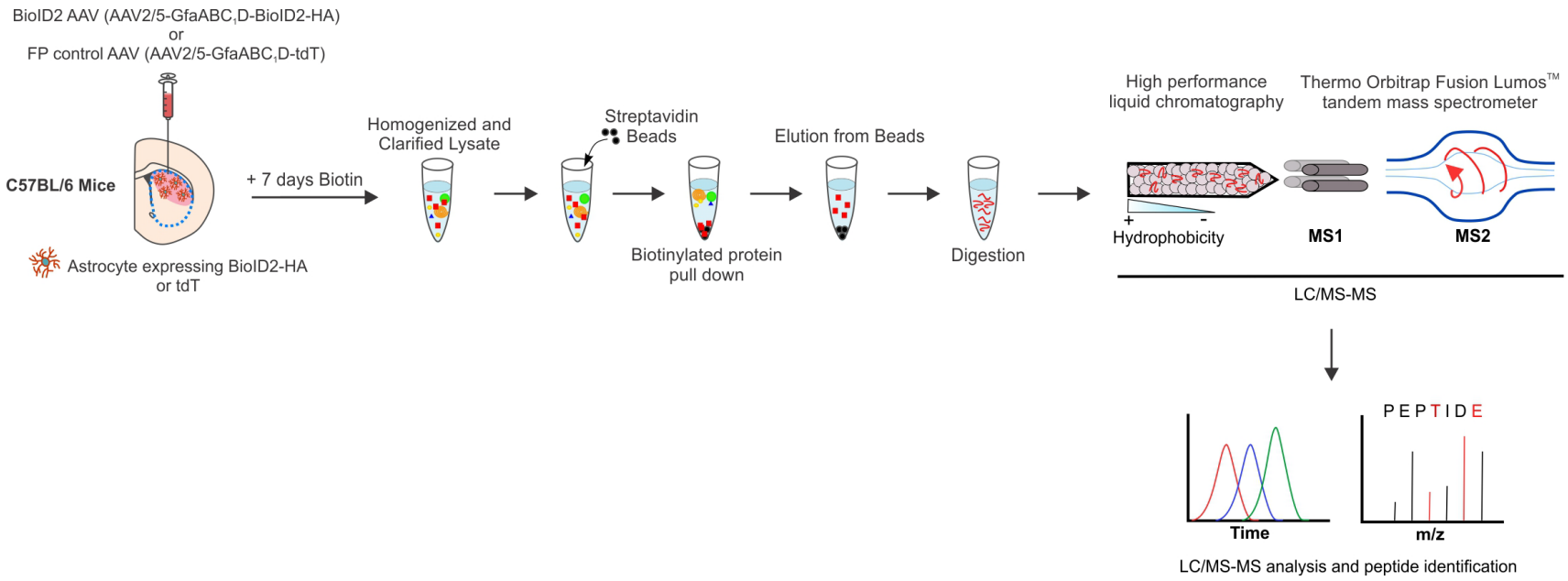

Figure 2.1: Biotinylated protein pulldown workflow. Schematic shows tissue processing steps to obtain biotinlylated proteins from striatal cells expressing BioID2. High performance liquid chromatography with tandem mass spectrometry (LC-MS/MS) is used to identify the biotinylated proteins

counting across multiple runs. MS2 spectra were searched using the ProLuCID algorithm against the Uniprot *Mus musculus* reference proteome (UP000000589) followed by filtering of peptide-to-spectrum matches (PSMs) by DTASelect with a "decoy" database-estimated false discovery rate (FDR) < 1%. In the second analysis, label-free quantification (LFQ) was carried out by integrating the total extracted ion chromatogram (XIC) at the MS1 level of parent peptide ions with a specific retention time and m/z through MaxQuant software as described previously (Cox et al,. 2008). To generate protein expression and quantification values, the LFQ intensity was used. The raw dataset generated by mass spectrometry has been deposited to the online database PRIDE with accession number PXD029257.

To account for variations in pulldown efficiency, all proteins and their LFQ values were robustly normalized to Pyruvate carboxylase (Uniprot ID Q05920). Downstream analysis was conducted only on proteins with non-zero LFQ values in 3 or more experimental BioID2 replicates. Data analysis for whole bulk tissue analyses was carried out identically except samples were normalized by median intensity.

Section 2.14: Statistical analysis of differentially expressed proteins

The LFQ intensity values were used for protein quantification across samples. Statistical analysis of differentially expressed proteins was done using Bioconductor package limma v 3.54. To generate a list of proteins with high confidence, all mitochondrial proteins including carboxylases and dehydrogenases were manually filtered as they are artifacts of endogenously biotinylated proteins. Proteins with $Log_2FC > 1$ and FDR < 0.05 versus GFP controls were considered putative hits and used for subsequent comparison between subcompartments and cell-types. Comparison between subcompartments and cell types was also performed with limma

utilizing the same thresholds ($Log_2FC > 1$ and FDR < 0.05). To generate a P-value for proteins that appeared in one group in > 3 replicates, but had a zero value in all the control experimental runs, the zero values were imputed with the lowest LFQ value in the entire dataset.

The gene ontology enrichment analysis for cellular compartments and biological function was performed using PANTHER overrepresentation test (GO ontology database released 2020-01-01) with FDR < 0.05 and with all *Mus musculus* genes used as reference and with STRING (string-db.org) with an FDR < 0.05 and with all *Mus musculus* genes used as a reference. Gene ontology pathway analysis for the astrocytic subcompartments was done with Enrichr (https://maayanlab.cloud/Enrichr/).

Section 2.15: Protein networks and protein-protein interaction analysis

Network figures were created using Cytoscape (v3.8) with nodes corresponding to the gene name for proteins identified in the proteomic analysis. A list of protein-protein interactions from published datasets was assembled using STRING. In all networks, node size is proportional to the fold enrichment over GFP control. To identify interactors of SAPAP3 protein, Significance Analysis of INTeractome (SAINTexpress; Choi et al,. 2013) was employed with a BFDR cutoff of 0.05.

Section 2.16: RNA-seq analysis of striatal astrocyte and neuron transcriptomes

RNA extraction from striatal astrocytes and neurons was performed as previously described (Nagai et al., 2019, Yu et al., 2020, Gangwani et al., 2023). Briefly, RiboTag AAV (AAV 2/5 $GfaABC_1D$ Rpl22-HA or AAV 1 hSyn1 Rpl22-HA) was injected into the dorsal striatum of adult C57BL/6NTac mice at 6 weeks of age. For RNA extraction of SAPAP3 KO mice, RiboTag AAV

was injected into the dorsal striatum at 4.5 months of age. Freshly microdissected striata were collected and individually homogenized. RNA was extracted from 10%-20% of cleared lysate as input sample, which contained RNA from all striatal cell types. The remaining lysate was incubated with mouse anti-HA antibody (1:250; Biolegend, 901514) with rotation for 4 hrs at 4°C followed by addition of IgG magnetic beads (Invitrogen Dynabeads #110.04D). The samples were left for overnight incubation at 4°C. The beads were then washed three times in high salt solution. RNA was purified from the immunoprecipitate and the corresponding input samples with the Qiagen RNAeasy kit (Qiagen 74034). RNA concentration and quality were assessed with Agilent 2100 Bioanalyzer. RNA samples with RNA integrity number (RIN) greater than 7 were used for multiplexted library preparation with the Nugen Ovation RNA-seq System V2. For each experiment, all samples were multiplexed into a single pool in order to avoid batch effects (Auer and Doerge, 2010). Sequencing was performed on Illumina NextSeq 4000 for 2 x 75 yielding at least 45 million reads per sample. Demultiplexing was performed with Illumina Bcl2fastq2 v 2.17 program. Reads were aligned to the mouse mm10 reference genome using the STAR spliced read aligner (Dobin et al., 2013). Approximately 70% of the reads mapped uniquely to the mouse genome and were used for downstream analyses. Differential gene expression analysis was performed on genes with FPKM > 0.01 in at least 4 samples per condition and $Log_2FC > 1$ or < -1 using Bioconductor package limmaVoom with the false discovery rate (FDR) threshold set at < 0.05. Combat was applied to remove batch effects. Differentially expressed genes (DEGs) that were more than two-fold higher in the IP samples than the input samples were designated as astrocyte or neuron enriched DEGs. RNAseq data has been deposited within the Gene Expression Omnibus (GEO) repository (https://www.ncbi.nlm.nih.gov/geo), accession ID# GSE184773.

Section 2.17: Behavioral tests of SAPAP3 WT and KO mice

Behavioral tests were performed during the light cycle between 10am and 2pm (Welch et al., 2007). Mice were assessed at 6 months of age or 5 months post-AAV microinjection. All experimental mice were transferred to the behavior testing room at least 30 min before testing to acclimatize to the environment and to reduce stress. Temperature and humidity of the experimental rooms were kept at 23 ± 2°C and 55 ± 5%, respectively. The brightness of the experimental room was kept dimly lit unless otherwise stated. Background noise (60-65 dB) was generated by a white noise box (San Diego Instruments).

Self-grooming behavior

The procedure of self-grooming behavior measurement was adapted from previously published work (Yu et al., 2018; Kalueff et al., 2007). The recording was conducted at 35 lux. Mice were placed individually into plastic cylinders (15 cm in diameter and 35 cm tall), and allowed to habituate for 20 minutes. Self-grooming behavior was recorded for 10 minutes. A timer was used to assess the cumulative time spent in self-grooming behavior which included paw licking, unilateral and bilateral strokes around the nose, mouth, and face, paw movement over the head and behind the ears, body fur licking, body scratching with hind paws, tail licking, and genital cleaning. The number of self-grooming bouts and rearing bouts was also counted. Separate grooming bouts were considered when the pause was more than 5 sec or if behaviors other than self-grooming occurred. The microstructure of self-grooming was not assessed.

Assessment of skin lesions

Mice were anesthetized in 5% isoflurane with 1% O_2 via a veterinary grade anesthesia ventilation system (VetEquip). Mice were placed on an opaque Plexiglass board and photos of their head and torso were taken bilaterally. Images were scaled with a ruler (Fine Science Tools) and the images were analyzed on ImageJ software. Measurements were all scaled to the ruler on ImageJ.

Open field test

The open field chamber was illuminated at 35 lux. The open field chamber consisted of a square arena (28 cm x 28 cm) enclosed by walls made of opaque white Plexiglass (19 cm tall). The periphery of the arena was defined as the area within 2.5 cm adjacent to the walls of the chamber and the center of the arena was defined as the area 2.5 cm away from the chamber walls. Each mouse was placed into an open field box to habituate for 20 min. Locomotor activity was then recorded for 30 minutes using a camera (Logitech) located right above the open field chamber. Anymaze video analysis software was used to quantify time spent in center, total distance traveled, and speed.

Elevated plus maze

All four arms of the elevated plus maze were illuminated at 25 lux. The elevated plus maze consisted of arms that were 30 x 7 cm with closed arm walls with a height of 20 cm. The maze was elevated 65cm above floor level and was placed in the center of the room away from other stimuli. Mice were placed in the center of the maze facing an open arm. Mice were recorded for 10 minutes using a camera (Logitech) located above the maze. Anymaze video analysis software was used to quantify time spent in open arms and percent time spent in open arms.

Section 2.18: Quantification and statistical analyses

Data from every experiment represent at least four replicates. All statistical tests, unless otherwise stated, were run in OriginPro 2018 and GraphPad InStat3. Data are presented as mean ± SEM along with the individual data points. The results of statistical comparisons, *n* numbers, and exact *p*-values are shown in the figure panels or in the figure legends along with the average data. *N* is defined as the number of cells or mice on a case-by-case basis throughout the manuscript. We determined whether each set of data was normally distributed or not by using GraphPad Instat3 and OriginPro 2018. If the data were normally distributed we employed parametric tests, while if they were not normally distributed we used parametric tests. Paired and unpaired Student's two-tailed t tests (as appropriate) and two tailed Mann-Whitney tests were used for most statistical analyses with significance declared at $p < 0.05$, but stated in each case with a precise *P* value. When *P* values were greater than 0.05, it is stated as non-significant (n.s.). All proteomic and transcriptomic analyses used a statistical FDR value < 0.05 unless otherwise stated. All mice were assigned to particular experimental groups at random. No data points were excluded from any experiment.

CHAPTER 3: ASTROCYTE AND NEURON PROTEOMES

Section 3.1: Introduction

In the central nervous system there exists various cell types. Neurons are the electrical conduits of the brain and are responsible for sending, receiving, and transmitting electrochemical signals throughout the entirety of the central nervous system. Glial cells are the cells which are not neurons and are responsible for all other aspects of brain function such as keeping the neuron axons ensheathed with myelin for fast electrical signal propagation (oligodendrocytes), immune surveillance (microglia), or maintenance of local circuitry (astrocytes). Although both astrocytes and neurons arise from the same developmental pathway during embryogenesis (Bayraktar et al., 2015) the fundamental difference between neurons and astrocytes are the core functions that each cell type conducts after differentiation. By definition, neurons are electrical conduits, and this function is attributed to the proteins related to electrical signaling it expresses. Thus, if this is the case, electrically "silent" cells such as astrocytes should express very little to no proteins related to electrical conductivity such as voltage gated Na+ channels. While there is a logic to these assumptions, there has been no direct systematic measure of the difference between astrocytes and neurons at the proteomic level in a defined circuit. Furthermore, while invaluable extensive databases have curated transcriptomic profiles of astrocytes and neurons with deep RNA-sequencing and single-cell sequencing (Zhang et al., 2014; Saunders et al., 2018; Tian et al., 2019) these profiles have not been compared to each other nor have they been compared to the resultant protein profiles of astrocytes or neurons.

In order to assess the core molecular distinctions and similarities between astrocytes and neurons at the protein level and *in vivo*, we developed a proximity labeling based approach to label and capture proteins in both cell types, selectively. To do this, we used genetically targeted biotin

ligase, BioID2. BioID2 is a small biotin ligase isolated from the organism *Aquifex aeolicus* with a humanized modification at the biotin catalytic domain (R40G) for more promiscuous biotinylation activity (Kim et al., 2016). In the presence of exogenous biotin, magnesium (Mg^{2+}), and ATP as an energy source, BioID2 catalyzes the fusion of biotin with a free lysine of a proximal protein (**Figure 3.1**), thus enabling the protein to be labeled and subsequently visualized or isolated with streptavidin, a strong biotin binder.

To selectively target either astrocytes or neurons in the striatum we utilized cell-selective AAVs. We utilized the human Synapsin-1 (*hSyn1)* gene promoter (Kügler et al., 2003) in combination with the AAV 1 serotype to selectively target neurons in the striatum. For astrocytes, we used the astrocyte minimal promoter, *GfaABC$_1$D* (Lee et al., 2008) in combination with the AAV 2/5 serotype to selectively target astrocytes in the striatum. In all constructs, BioID2 contained a hemagglutinin (HA) tag at the C-terminus in order to be able to visualize its specificity and localization in the striatum with antibody staining. In this way, we created Neuro BioID2 (*hSyn1* BioID2-BioID2-HA) and Astro BioID2 (*GfaABC$_1$D* BioID2-BioID2-HA) which localized freely diffusible, cytosolic BioID2 to either neurons or astrocytes (**Figure 3.2**).

Additionally, to take advantage of the nanoscale range of BioID2 biotinylation, we created constructs that targeted BioID2 to the plasma membrane. The Src family of protein tyrosine kinases contain an N-terminal motif that specifies modifications necessary for precise intracellular localization to the plasma membrane. This motif, termed the Lck peptide, sequence allows direct protein targeting to the plasma membrane (Zlatkine et al., 1997). Using the Lck peptide, we created Neuro Lck-BioID2 (*hSyn1* Lck-BioID2-BioID2-HA) and Astro Lck-BioID2 (*GfaABC$_1$D* Lck-BioID2-BioID2-HA) for labeling and capture of proteins at the plasma membrane of these cells (**Figure 3.2**). With the development of these cell and subcompartment-specific constructs, our

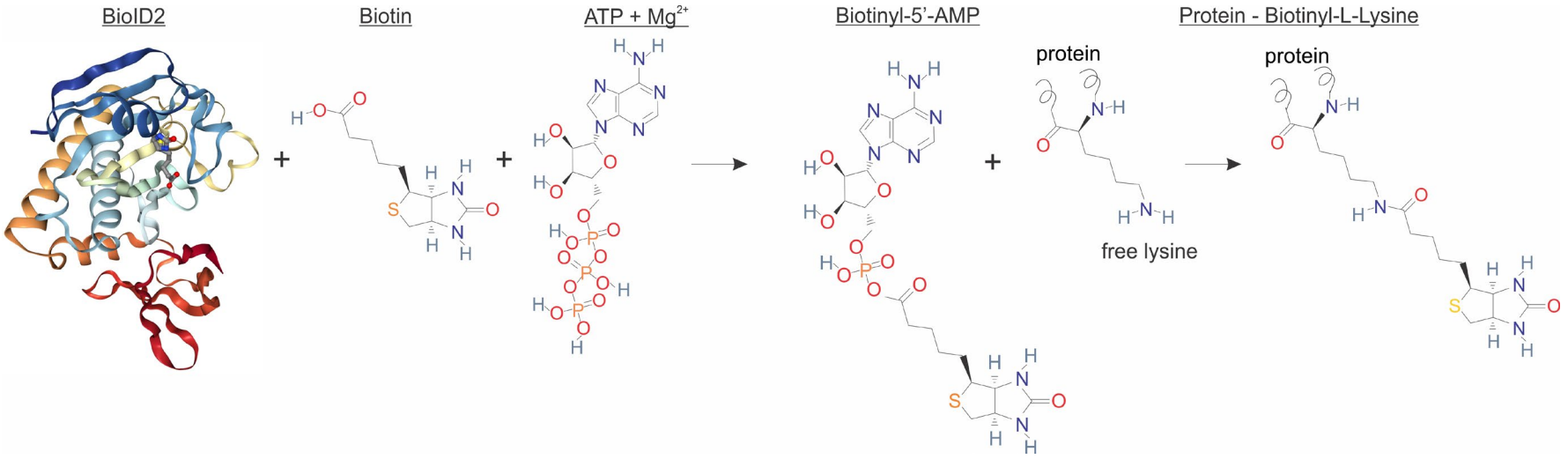

Figure 3.1: Chemical reaction catalyzed by BioID2. Schematic of the BioID2 biotinylation reaction in the presence of biotin, Mg^{2+}, and ATP. Free lysine residues on a protein are biotinylated.

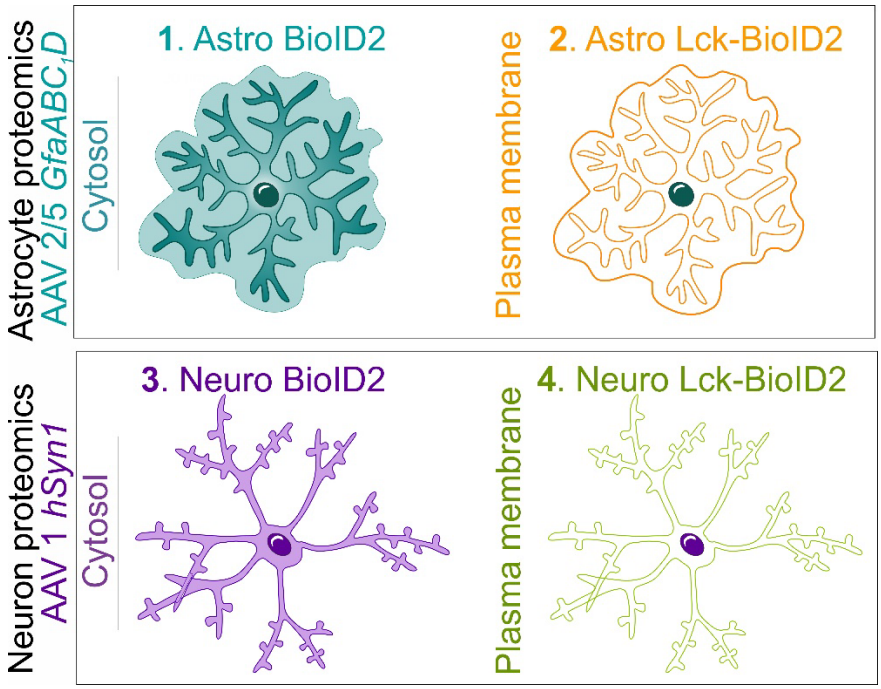

Figure 3.2: Cell and subcompartment specific AAVs. Cartoon illustrating four adeno-associated virus (AAV) constructs for cell and subcompartment specific targeting of BioID2 in astrocytes and neurons.

studies provide a basis for functional differences between astrocytes and neurons with some surprising similarities.

Section 3.2: Validation of BioID2 subcompartment localization with HEK293 cells

In order to assess whether BioID2 could be targeted to the plasma membrane in mammalian cells, we used HEK293T kidney cells to explore BioID2 expression before applying BioID2 *in vivo*. We made to expression construct plasmids using pcDNA3.1 as the vector for mammalian expression and used the cytomegalovirus (CMV) promoter for expression in mammalian cells. This resulted in two constructs: pcDNA3.1 CMV BioID2-BioID2-HA for cytosolic BioID2 expression and pcDNA3.1 CMV Lck-BioID2-BioID2-HA for plasma membrane expression.

When cells were transfected with either of these plasmids, we observed that both BioID2 constructs reliably expressed in mammalian cells (**Figure 3.3**) and in the case of plasma membrane Lck-BioID2, BioID2 was highly expressed at the plasma membrane. Interestingly, cytosolic BioID2 was also shown to be expressed in the nucleus of the cell, indicating that freely diffusing BioID2 is small enough to cross the nuclear envelope. This nuclear expression was not seen in the case of Lck-BioID2.

Upon addition of exogenous biotin at a concentration of 5 mM, we detected a strong biotinylated protein signal when the cells were probed with streptavidin. This signal was only detected upon addition of exogenous biotin as compared with media-only controls (**Figure 3.3**). This result indicated that (1) BioID2 does not sequester endogenous biotin and (2) biotinylation of proteins by BioID2 is temporally specific and can be controlled by the addition of exogenous biotin. Furthermore, biotinylated proteins were highly co-localized with the BioID2-HA expression signal indicating that indeed, BioID2 is proximity dependent and labels proteins within

10nm. This was exceptionally clear for plasma membrane targeted Lck-BioID2 as most of the biotinylated protein signal was found at the plasma membrane of the HEK cell.

Because localized expression of BioID2 and subsequent protein biotinylation was possible in mammalian HEK293 cells, we then conducted a western blot for biotinylated proteins in these cells in order to assess the range of biotinylation by BioID2. Upon gel separation of proteins and western blotting with Streptavidin conjugated-horseradish peroxidase (HRP) staining, we found that a range of proteins across the entire molecular weight spectra was biotinylated by both cytosolic and plasma membrane localized BioID2 constructs (**Figure 3.4**). Furthermore, in concordance with the immunocytochemistry data, the signal of this range of biotinylated proteins was significantly higher upon addition of exogenous biotin (**Figure 3.4**) Since biotinylated proteins were detected at the western blot level and these proteins ranged across molecular weights, the results indicated that these proteins could be identified with mass spectrometry. Therefore, subcompartment targeted BioID2 can be used to label and identify proteins in mammalian cells.

When conducting these BioID2 western blot analyses, it is important to note two strong biotinylated protein bands at 130 kDa and 250 kDa in both non-biotin and biotin conditions. These are endogenously biotinylated mitochondrial proteins that exist within a cell (Grant et al., 2019) and should be interpreted with this in mind.

Section 3.3: *In vivo* expression of BioID2 in astrocytes and neurons

Before dissecting the proteomes of astrocytes and neurons in both the cytosol and plasma membrane, we first validated the expression and biotinylation activity of Astro BioID2, Astro Lck-BioID2, Neuro BioID2, and Neuro Lck-BioID2 in the striatum with immunohistochemistry and western blots. To do this, we microinjected each AAV construct unilaterally into the dorsal

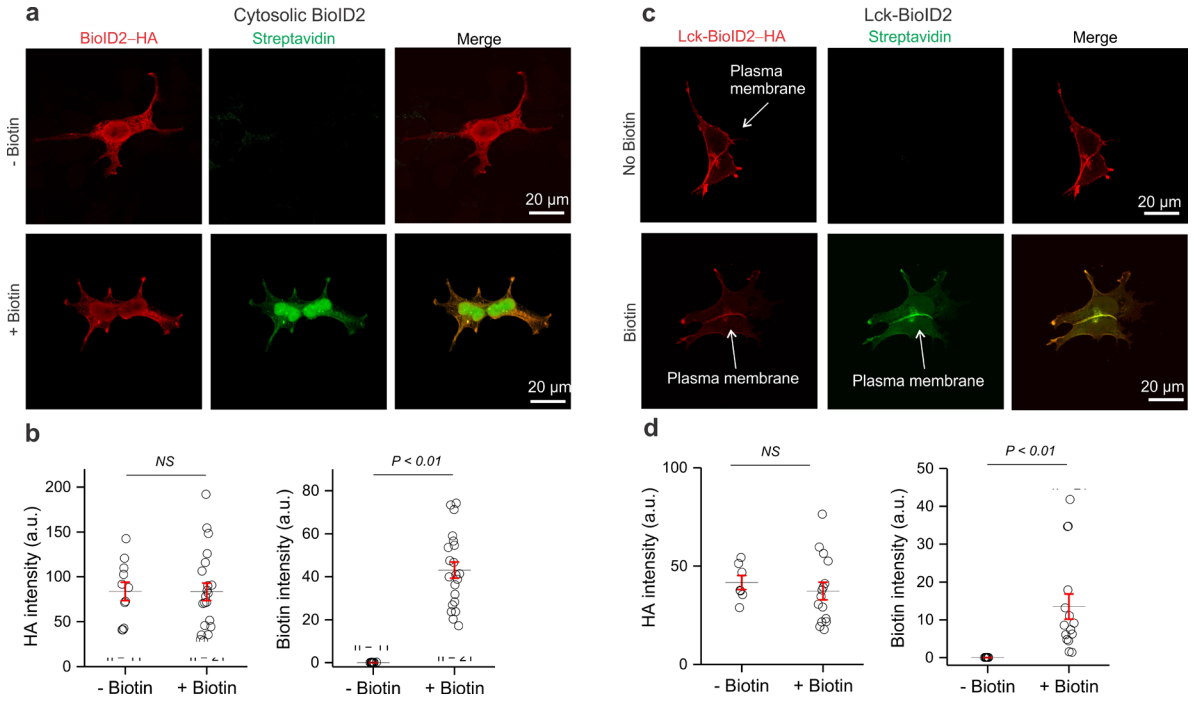

Figure 3.3: Compartment specific BioID2 validation in HEK293T cells. a. Representative images of HEK293T cells transfected with the cytosolic BioID2 construct and then treated with either biotin in PBS or solely PBS (no biotin). The cells were immunostained with both anti-HA antibody (red) and a fluorophore conjugated streptavidin probe (green). **b.** Quantification of HA intensity and biotin intensity (arbitrary units) in cells transfected with cytosolic BioID2 (n = 11 cells in the no biotin group and n = 21 cells in the biotin group; Unpaired T-test) **c.** Representative images of HEK293T cells transfected with the plasma membarane Lck-BioID2 construct and then treated with either biotin in PBS or solely PBS (no biotin). The cells were immunostained with both anti-HA antibody (red) and a fluorophore conjugated streptavidin probe (green). Arrows show the plasma membrane localization. **f.** Quantification of HA intensity and biotin intensity (arbitrary units) in cells transfected with Lck-BioID2 (n = 7 cells in the no biotin group and n = 21 cells in the biotin group; Unpaired T-test).

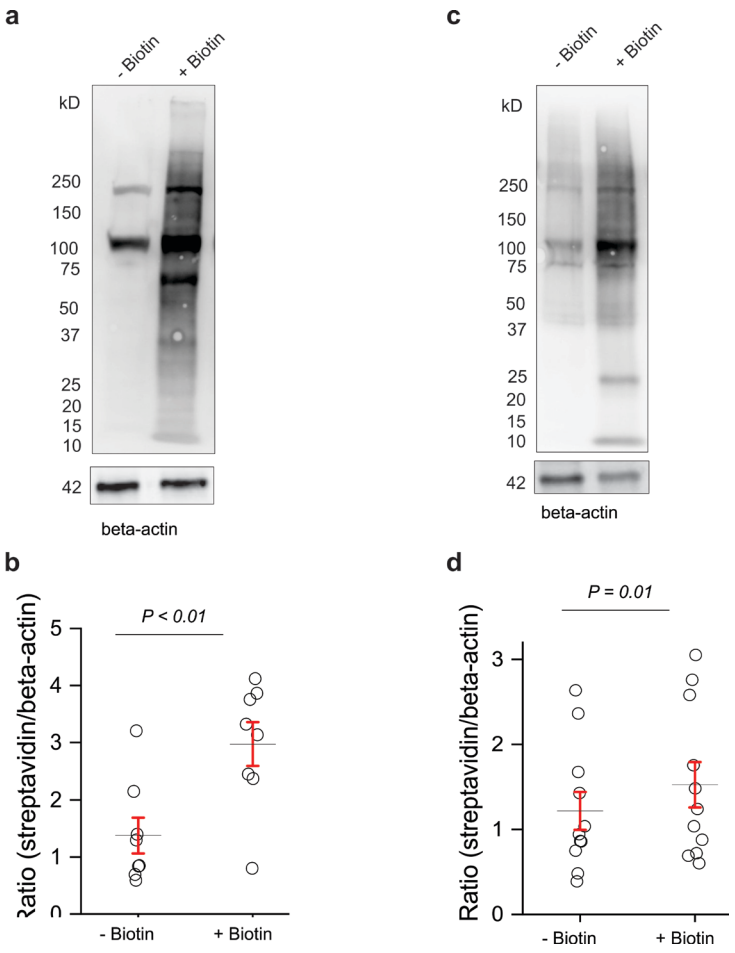

Figure 3.4: Biotinylation validation of BioID2 in HEK293T cells. a. Representative western blot of HEK293T cells transfected with cytosolic BioID2 and treated with either biotin in PBS or solely PBS (no biotin). **b**. Western blot quantification of HEK293T cells transfected with cytosolic BioID2 and treated with either biotin in PBS or solely PBS (no biotin). Graph depicting the streptavidin signal intensity divided by the β-actin signal intensity for each data point. Black horizontal line depicts the mean. (n = 8 batches of coverslips; Unpaired T-test). **c**. Representative western blot of HEK293T cells transfected with Lck-BioID2 and treated with either biotin in PBS or solely PBS (no biotin). **d**. Western blot quantification of HEK293T cells transfected with Lck-BioID2 and treated with either biotin in PBS or solely PBS (no biotin). Graph depicting the streptavidin signal intensity divided by the β-actin signal intensity for each data point. Black horizontal line depicts the mean. (n =11 batches of coverslips; Unpaired T-test).

striatum. After three weeks of expression, a subcutaneous injection of exogenous biotin was given to the mice for seven days. Sixteen hours after the last injection, the mouse brain tissues were processed for either immunohistochemistry or western blot.

With immunochemistry we found that BioID2-HA expression was localized to the striatum, the intended expression site. When the tissue was probed with streptavidin, the biotinylated protein signal was the highest and restricted to the area where BioID2 was expressed, thus indicating that BioID2 expression and subsequent protein biotinylation can be restricted to discrete brain areas and defined circuits (**Figure 3.5**). When assessing the biotinylated proteins with western blotting and the streptavidin-HRP probe, all four cell-specific and compartment specific constructs displayed a significant increase in biotinylation signal across the molecular weight spectra when compared to the contralateral side (**Figure 3.5**). Therefore, BioID2 can be expressed and is functional in brain cells, *in vivo.*

We next assessed whether our *in vivo* BioID2 constructs were indeed cell-selective and subcompartment-specific as intended. To do this we conducted immunohistochemistry of BioID2 and the biotinylated proteins along with cell markers NeuN and DARPP-32 medium spiny neurons (MSNs) in the striatum and S100β for striatal astrocytes. Our results show that both Astro BioID2 and Astro Lck-BioID2 were expressed in 85% of S100β+ astrocytes and in about 1% of neurons (**Figure 3.6**). Additionally, visualization of Astro BioID2 and its corresponding biotinylated protein signal looked both more somatic and end-foot vasculature-like than Astro Lck-BioID2 which displayed a cloud-like bushy signal reminiscent of the astrocyte fine processes. This indicated that both Astro BioID2 and Astro Lck-BioID2 were selectively expressed in striatal astrocytes and biotinylated proteins at their corresponding subcompartments *in vivo*.

Similar immunohistochemistry experiments were conducted with the Neuro BioID2 and Neuro Lck-BioID2 constructs. In this case about 85% of MSNs expressed Neuro BioID2 or Neuro Lck-BioID2, whereas no S100β+ astrocytes expressed either construct. Visualization of Neuro BioID2 and its corresponding biotinylated protein signal showed discrete neuronal somata labeling while Neuro Lck-BioID2 displayed a dense and dendritic-like expression. Therefore, both Neuro BioID2 and Neuro Lck-BioID2 were selectively expressed in striatal neurons and at their corresponding subcompartments (**Figure 3.7**).

Section 3.4: Mass spectrometry based identification and analysis of biotinylated proteins

Western blot analyses of striata expressing cell and subcompartment specific BioID2 demonstrated that a range of proteins are biotinylated *in vivo*. Thus, we next sought to isolate and identify these biotinylated proteins using mass spectrometry based methods. We injected a set of mice with either cell type- or compartment specific BioID2 AAVs bilaterally into the striatum. We also injected another set of mice with a cognate cell-specific GFP AAV control such as cytosolic tdTomato or plasma membrane tethered Lck-GFP as a control. After three weeks of expression, all mice received the same subcutaneous injection of 24 mg/kg of biotin for 7 days, after which the mouse striata were processed for purification of biotinylated proteins.

After purification of the proteins with streptavidin conjugated beads, the proteins were trypsinized to yield peptide fragments for downstream mass spectrometry analyses. We used high performance liquid chromatography coupled with tandem mass spectrometry to assess the identity and relative abundance of these cell specific peptide fragments. Using the Uniprot *mus musculus* database as a reference, a total of 113,421 unique peptides were run across all cell-specific samples including the controls. We identified 3525 proteins in the combined Neuro BioID2 and Neuro GFP

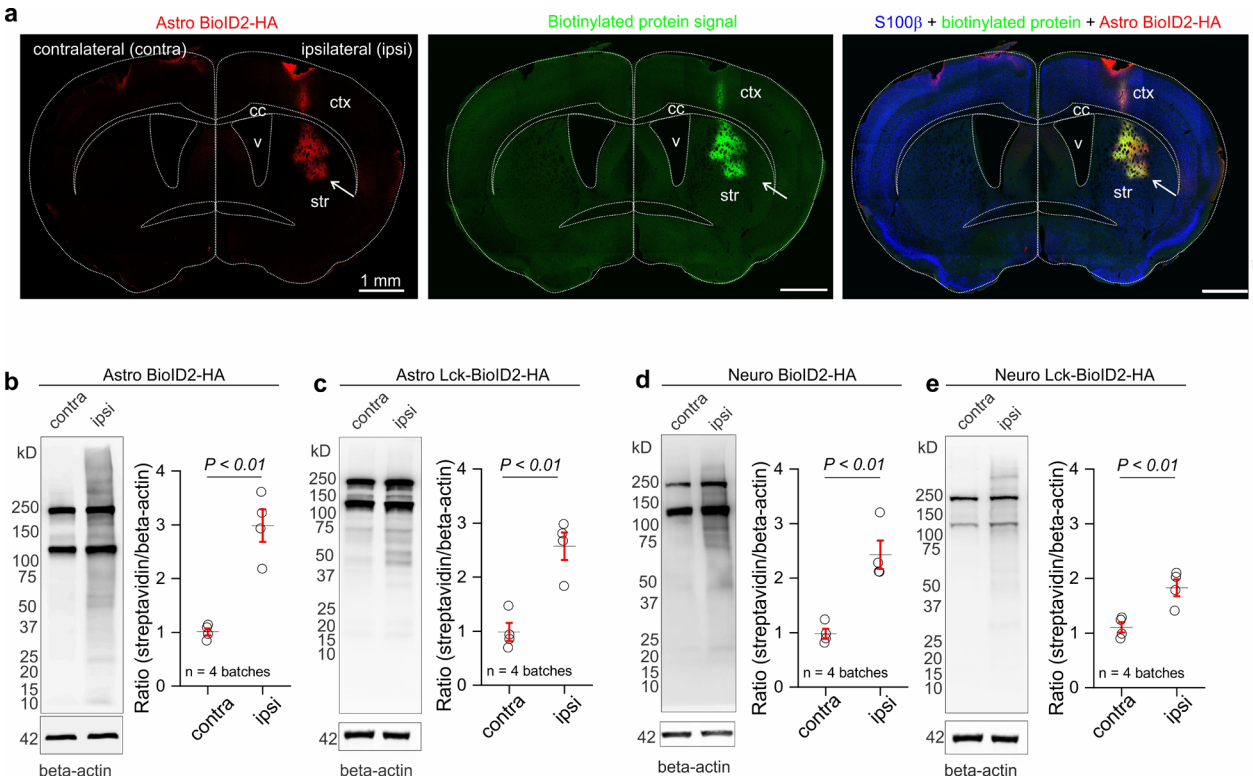

Figure 3.5: *in vivo* expression of BioID2 in mouse striatum. a. Coronal section of mouse brain microinjected in the dorsal striatum with AAV 2/5 *GfABC1D* BioID2-HA (Astro BioID2) and injected subcutaneously with biotin for 7 days. **b.** Western blot analysis of brain unilaterally microinjected with Astro BioID2. Dark bands at 130 kD and 250 kD show the endogenously biotinylated mitochondrial proteins, Pyruvate carboxylase and acetyl-CoA carboxylase. Graph depicting the streptavidin signal intensity divided by the β-actin signal intensity for each data point. Black horizontal line depicts the mean. (n = 4 mice; Paired T-test) **c.** As in **b**, but for membrane localized Astro Lck-BioID2. (n = 4 mice; Paired T-test) **d.** As in **b**, but for AAV1 hSyn1 BioID2-HA (Neuro BioID2). . (n = 4 mice; Paired T-test). **e.** As in **b**, but for plasma membrane localized Neuro Lck-BioID2 (n = 4 mice; Paired T-test).

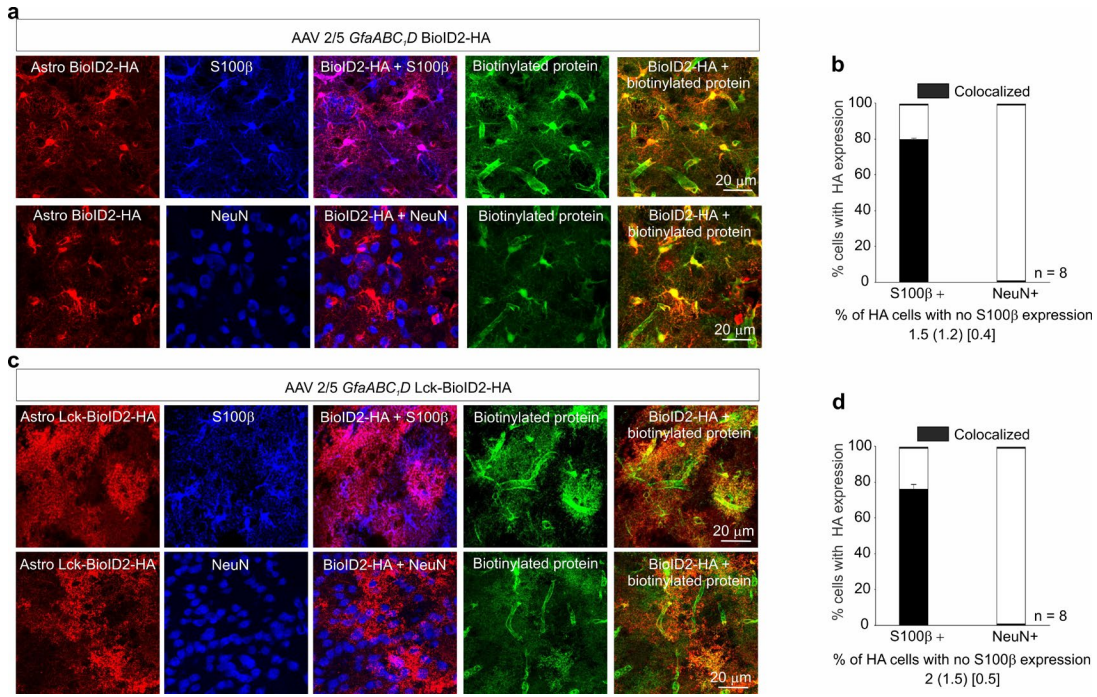

Figure 3.6: Cell-selectivity of BioID2 in astrocytes. a. Representative images of mouse striatum injected with astrocyte-specific cytosolic BioID2 and then treated with biotin for 7 days. Top panel shows the immunostaining pattern with S100β as an astrocyte cell marker and bottom panel shows the immunostaining pattern with NeuN as a neuronal cell marker. **b.** Bar graphs depicting the percent of S100β positive or NeuN positive cells with HA expression in a 40x magnification field of view. Black portion of the bar graphs show the percent co-localization. (n = 8 fields of view at 40x magnification from 4 mice) **c.** Representative images of mouse striatum injected with astrocyte-specific plasma membrane Lck-BioID2 and then treated with biotin for 7 days. Top panel shows the immunostaining pattern with S100β as an astrocyte cell marker and bottom panel shows the immunostaining pattern with NeuN as a neuronal cell marker. **d.** Bar graphs depicting the percent of S100β positive or NeuN positive cells with HA expression in a 40x magnification field of view. Black portion of the bar graphs show the percent co-localization. (n = 8 fields of view at 40x magnification from 4 mice).

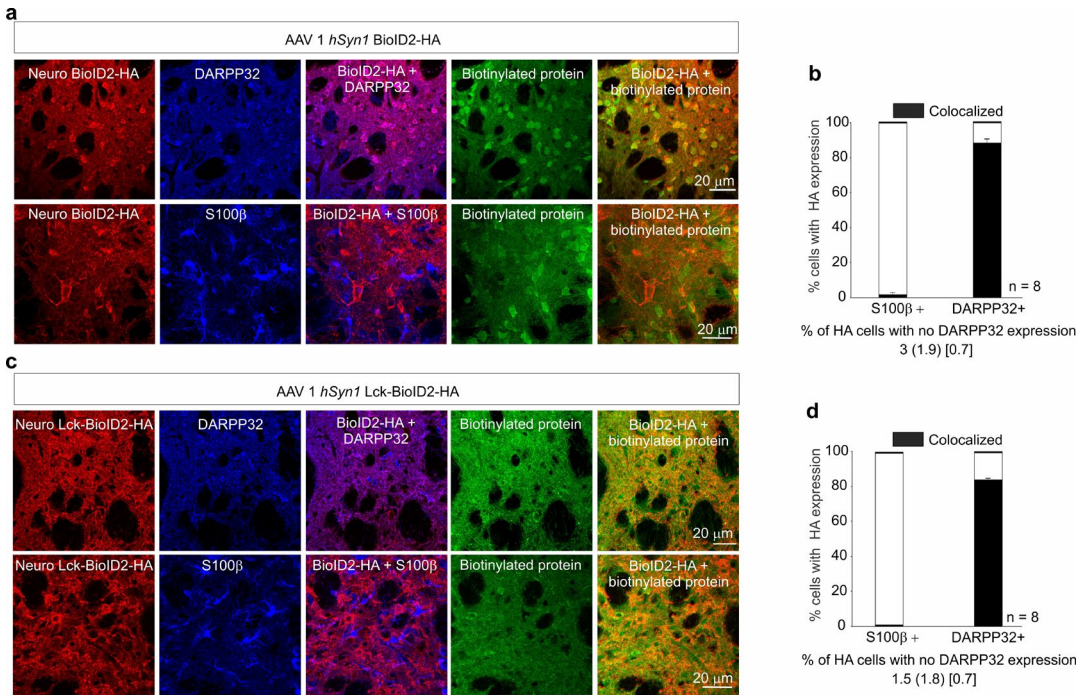

Figure 3.7: Cell-selectivity of BioID2 in neurons. a. Representative images of mouse striatum injected with neuron-specific cytosolic BioID2 and then treated with biotin for 7 days. Top panel shows the immunostaining pattern with NeuN as a neuronal cell marker and bottom panel shows the immunostaining pattern with S100 β as an astrocyte cell marker. **b.** Bar graphs depicting the percent of S100β positive or NeuN positive cells with HA expression in a 40x magnification field of view. Black portion of the bar graphs show the percent co-localization. (n = 8 fields of view at 40x magnification from 4 mice) **c.** Representative images of mouse striatum injected with neuron-specific plasma membrane Lck-BioID2 and then treated with biotin for 7 days. Top panel shows the immunostaining pattern with NeuN as a neuronal cell marker and bottom panel shows the immunostaining pattern with S100β as an astrocytic cell marker. **d.** Bar graphs depicting the percent of S100β positive or NeuN positive cells with HA expression in a 40x magnification field of view. Black portion of the bar graphs show the percent co-localization. (n = 8 fields of view at 40x magnification from 4 mice).

identification list, 3603 proteins in the combined Neuro Lck-BioID2 and Neuro Lck-GFP identification list, 3509 proteins in the combined Astro BioID2 and Astro tdTomato identification list, and 3754 proteins in the combined Astro Lck-BioID2 and Astro Lck-GFP identification list. The proteins identified here represent proteins that were identified in at least one technical run. A total of 4 technical and biological runs consisting of 8 mouse striata each were run to generate these lists.

In order to identify as many proteins as possible and keep sample quantity intact, we did not use any labeling methods such as tandem mass tagging (TMT) or isotope dimethyl labeling for protein quantitation. Instead we utilized label free quantification (LFQ) which is based on peptide spectral counting and the precursor (MS1) signal to generate a relative quantification termed the LFQ intensity (Griffin et al., 2010). If the LFQ intensity was 0, we deemed this protein non-existent in the sample. We normalized the LFQ values among technical runs by using the LFQ values of the endogenously biotinylated 130 kDa mitochondrial protein, Pyruvate carboxylase, which appears robustly at similar abundance levels in all samples, including in the fluorescent protein controls.

After normalization with Pyruvate carboxylase, we conducted a dimensionality reducing analysis using principal component analysis (PCA) of all the proteins in order to visualize how each cell-specific BioID2 and their corresponding fluorescent protein control experiment related to each other. From this analysis, we found that principal component 1 explained about 45% of the variation of each construct sample and that the fluorescent protein controls separated far on this plane from all BioID2 experiments (**Figure 3.8**). We then manually removed all mitochondrial associated proteins as these can all be transiently and endogenously biotinylated. Additionally we removed proteins labeled by Uniprot as known contaminants such as mouse keratins. This left us

with 1509 proteins in the combined Neuro BioID2 and Neuro GFP identification list, 2022 proteins in the combined Neuro Lck-BioID2 and Neuro Lck-GFP identification list, 2091 proteins in the combined Astro BioID2 and Astro tdTomato identification list, and 1071 proteins in the combined Astro Lck-BioID2 and Astro Lck-GFP identification list. Finally, using the LFQ values and using R statistical package, limma, we chose high confidence proteins that had a larger than 2-fold LFQ intensity over the fluorescent protein controls and had a Bonferroni false discovery rate (adjusted p-value) of less than 0.05. This analysis yielded 590 proteins in the Neuro BioID2 identification list, 1942 proteins in the Neuro Lck-BioID2 identification list, 513 proteins in the Astro BioID2 identification list, and 385 proteins in the Astro Lck-BioID2 identification list. These values and filters are found in **Table 3.1 and supplementary excel file 1**.

Section 3.5: Characterization of astrocyte and neuronal proteomes

Mass spectrometry and subsequent statistical analyses allowed us to generate a high confidence list of proteins for each cell-specific and compartment specific BioID2 experiments. These high confidence lists represent proteins that were biotinylated by BioID2 in a cell-specific manner *in vivo*. To assess this cell-specificity we asked whether proteins that are considered cell-enriched proteins for multiple brain cell types were present or absent from each cell- and subcompartment specific experiment. For neurons we chose NeuN (gene: *Rbfox3*), Nefl (gene: *Nefl*), β-III tubulin (gene: *Tubb3*), and Enolase-2 (gene: *Eno2*). For astrocytes we chose Aldh1*l*1 (gene: *Aldh1l1*), Kir4.1 (gene: *Kcnj10*), Glt-1 (gene: *Slc1a2*) and Cx-43 (gene: *Gja1*). For oligodendrocytes we chose MOG (gene: *Mog)*, Sox10 (gene: *Sox10*), MOBP (gene: *Mobp*), and MBP (gene: *Mbp*). For oligodendrocyte progenitor cells we chose Olig2 (gene: *Olig2*), PDGFR-α, (gene: *Pdgfra*) and Cacng4 (gene: *Cacng4*). For microglia we chose CD-68 (gene: *Cd68*), Ptprc (gene: *Ptprc*), CCL-

3 (gene: *Ccl3*), and TNF (gene: *Tnf*). When we searched our high confidence protein lists for these cell-enriched markers we found that the all four neuron enriched markers were found within those proteins biotinylated by Neuro BioID2. Interestingly, only three of the four neuronal markers were found within proteins biotinylated with Neuro Lck-BioID2 (**Figure 3.9**). Nuclear protein, NeuN, was not found in the Neuro Lck-BioID2 protein list indicating that Neuro Lck-BioID2 only biotinylates proteins that associate with the plasma membrane at some point and that sub-compartment specificity is attained by genetically targeted BioID2. No cell-enriched markers for any other cell types were biotinylated by either Neuro BioID2 or Neuro Lck-BioID2 (**Figure 3.9**). We conducted these same analyses for both Astro BioID2 and Astro Lck-BioID2 and found that both constructs biotinylated all four astrocyte enriched proteins and no other cell type enriched marker proteins (**Figure 3.9**). Taken together, these analyses support the immunohistochemical data that our BioID2 AAV constructs are both cell-selective and sub-compartment targeted.

Comparison of our high confidence proteins to each other revealed common proteins, shared proteins, and proteins unique to each BioID2 construct. There were 5 proteins shared by all cell- and subcompartment specific constructs. 350 proteins were unique to Neuro Lck-BioID2, 109 proteins were unique to Neuro BioID2, 82 proteins were unique to Astro Lck-BioID2, and 78 proteins were unique to Astro BioID2. 436 proteins were shared by both neuronal BioID2 constructs, while 147 were shared by both astrocyte BioID2 constructs (**Figure 3.10**).

To make sense of the differences between astrocytes and neurons we compared both the astrocyte and neuronal cytosolic proteomes to each other and we also compared the astrocyte and neuronal plasma membrane proteomes in a similar manner. Using the LFQ intensity and R statistical analysis, limma, we quantitatively compared the high confidence proteins of the neuron and astrocyte cytosol to each other. We found 372 unique neuronal cytosol proteins and 302

astrocyte cytosol unique proteins. Among these neuronal unique proteins were proteins known to be in neurons such as Bassoon (*Bsn*) and neurofilament protein *(Nefl*). Proteins uniquely in astrocytes were proteins such as Glt-1 (*Slc1a2*) and the astrocytic sodium potassium pump (*Atp1a2*). Additionally, there were 208 shared proteins among the cytosol of both cell types. Of these, 62 were significantly enriched in neurons and 31 were enriched in astrocytes. Intriguingly, there were 115 proteins that were not enriched in either cell-type and are thus equivalently expressed in both astrocytes and neurons (**Figure 3.11**). Henceforth, the 372 neuronal unique proteins with the addition of the 62 enriched proteins when compared to the astrocyte cytosol comprise the high confidence striatal neuron cytosolic proteome for a total of 434 proteins. Likewise, the 302 astrocyte unique proteins with the addition of the 31 enriched proteins when compared to the neuron cytosol comprise the high confidence striatal astrocyte cytosolic proteome for a total of 332 proteins.

Similarly, we conducted the same analyses with the neuronal and astrocytic plasma membrane. When we quantitatively compared the proteins found at the neuron and astrocyte plasma membranes, we find 1640 were unique to the neuron plasma membrane including microtubule associated proteins (Map1a, Map1b) and post-synaptic associated proteins such as Ank-2 (*Ank2*). 246 proteins were unique to the astrocyte plasma membrane including Glt-1 (*Slc1a2*) and a bicarbonate transporter (*Slc4a4*). 262 proteins were shared among both cell types' plasma membranes, 31 of which were enriched in neurons and 64 were enriched in astrocytes. Again, there were 167 proteins that were found equivalently in both cells' plasma membranes including a surprising post-synaptic associated protein, SAPAP3 (*Dlgap3*) (**Figure 3.11**). Thus, the 1640 neuron plasma membrane unique proteins with the addition of the 31 neuron enriched proteins comprise the striatal neuron plasma membrane proteome comprised of 1672 proteins. While the

striatal astrocyte plasma membrane proteome is comprised of 310 proteins composed of the 246 unique and 64 enriched proteins when compared to the neuronal plasma membrane.

Section 3.6: Functional assessments of striatal astrocyte and neuronal proteomes

We next sought the functional relevance of the astrocyte and neuronal proteomes we characterized. To do this, we used PANTHER (Protein Analysis Through Evolutionary Relationships) protein classification and gene ontology analysis to assess the functions of the proteins in each cell compartment. Analysis of the astrocyte cytosolic proteins revealed that at least 10% (33 proteins) were involved in lipid metabolism. Other functions of astrocyte cytosolic proteins including cell-cell signaling and cell localization. Astrocyte plasma membrane proteins showed similar functions in lipid metabolism as well as about 20% of proteins involved in actin based processes. Neuronal cytosolic proteins were involved in signaling and ion binding while neuronal plasma membrane proteins showed similar signaling receptor activity as well as synaptic signaling. These results indicate that at the protein level, astrocytes and neurons differ by the types of proteins they express and the activities of these proteins. Of note is that about 45% of proteins identified in each cell compartment were proteins related to biosynthetic pathways, which may reflect homeostatic mechanisms conserved in both astrocytes and neurons (**Figure 3.12**).

Synaptic signaling in neurons and lipid metabolism in astrocytes were the most divergent functional pathways between astrocytes and neurons. To assess more specific molecular pathways that give rise to these differences, we probed our protein datasets for proteins involved in Ca^{2+} dependent vesicle release proteins, a canonical neuronal pathway that is necessary for neurotransmission. There are 28 known proteins and their isoforms that are involved in Ca^{2+} dependent vesicle release. These include proteins in the Complexin family, proteins in the Syntaxin

family, Proteins in the SNAP receptor family, proteins in the Synaptotagmin family, and proteins in the Synaptobrevin family. The minimum requirement for functional Ca^{2+} dependent vesicle release is to express proteins in all these families together. When we assessed these protein families in our striatal neuron proteomes, we found that the combined neuronal cytosol and plasma membrane proteomes contained proteins necessary for functional Ca^{2+} dependent vesicle release including the vesicular GABA transporter (*Slc32a1*) which has been shown to be abundantly expressed in medium spiny neurons of the striatum. However, the astrocyte proteomes revealed that astrocytes do not express the minimum machinery necessary to conduct Ca^{2+} dependent vesicle release, with only three proteins identified: SNAP25 (*Snap25*), Syntaxin1b (*Stx1b*), and Synaptobrevin-2 (*Vamp2*) (**Figure 3.13**).

In this way, we also assessed proteins that are involved in lipid metabolism, a pathway necessary for maintenance of membranes. We identified 53 total lipid metabolism related proteins among all proteomes, and found that astrocytes express the majority of them. About 35 proteins related to lipid metabolism were detected in astrocytes, while only about 5 were detected in neurons thus showing that astrocytes may be the molecular drivers of lipid metabolism in the striatum (**Figure 3.13**).

Our data show that astrocytes and neurons differ fundamentally by distinct molecular pathways that give rise to distinct astrocyte and neuronal functions.

Section 3.7: Correlations between transcriptomic and proteomic data

Transcriptomic analysis of specific cell types has been a standard in biology since its inception in the early 21st century. While there exists a plethora of transcriptomic datasets, including some *in vivo*, available from both astrocytes and neurons, there has not been a comparison between these

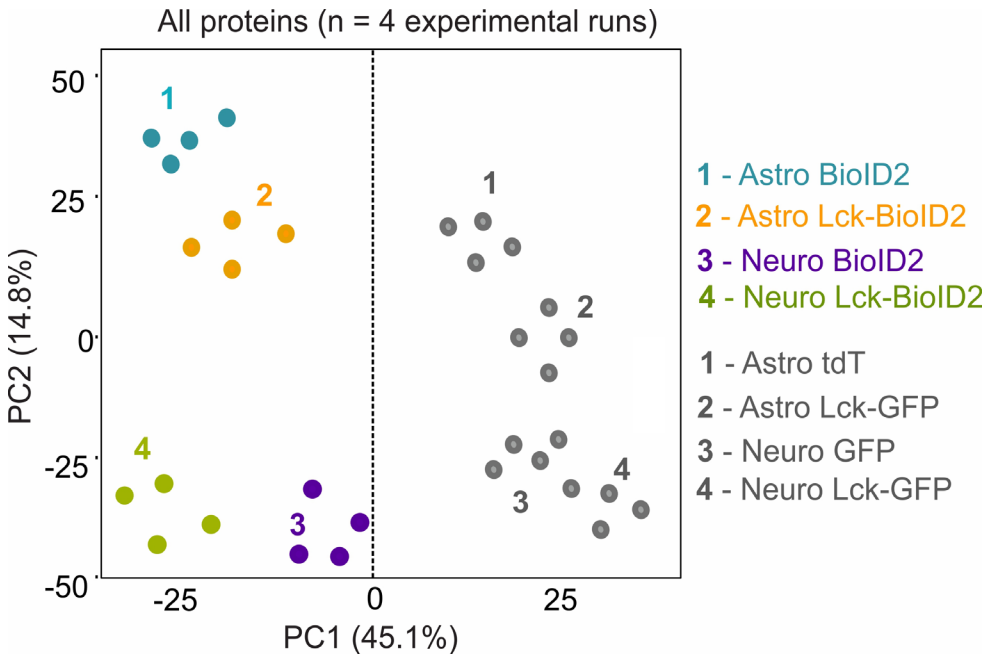

Figure 3.8: Principal component analysis of all proteins identified by mass spectrometry. PCA plot depicting all proteins identified by mass spectrometry across 4 experimental runs per construct.

Filters	Astro BioID2	Astro Lck-BioID2	Neuro BioID2	Neuro Lck-BioID2
# unique peptides	113,421			
#unique proteins	3,509	3,754	3,525	3,603
Proteins known to be contaminants and mitochondrial	2,091	1,071	1,509	2,022
# proteins with 2-fold over GFP controls and FDR <0.05	513	385	590	1,942

Table 3.1: BioID2 proteomics analysis filters. List of filters and resultant number of proteins used to generate high confidence lists for both astrocyte and neuron specific proteomics.

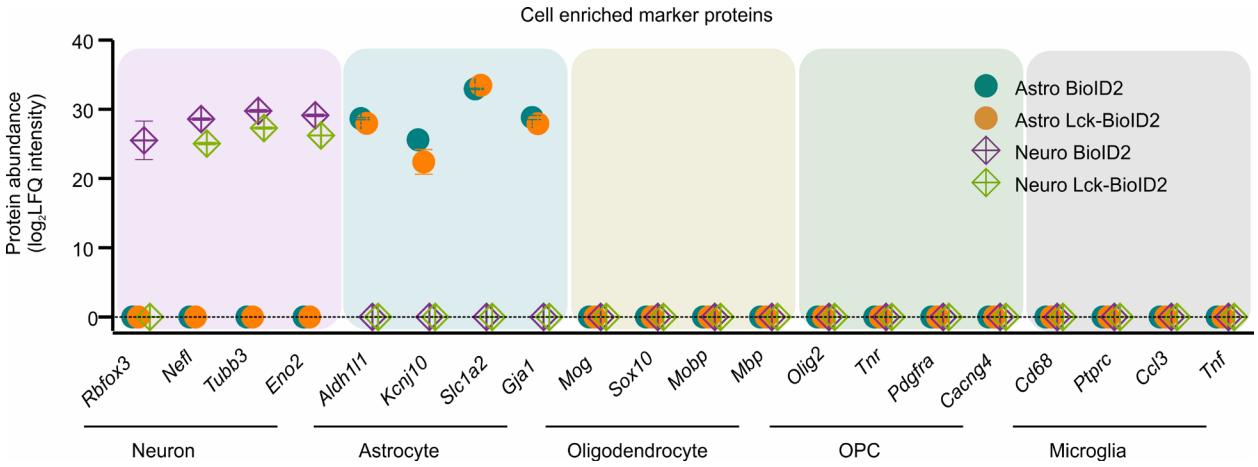

Figure 3.9: Cell-selectivity assessment of identified proteins. Expression levels (in label-free quantification intensity, LFQ intensity) of cell-specific markers identified in each BioID2 construct (n = 4 experimental runs). Proteins in this analysis represent proteins that were enriched (Log2FC > 1 versus GFP controls).

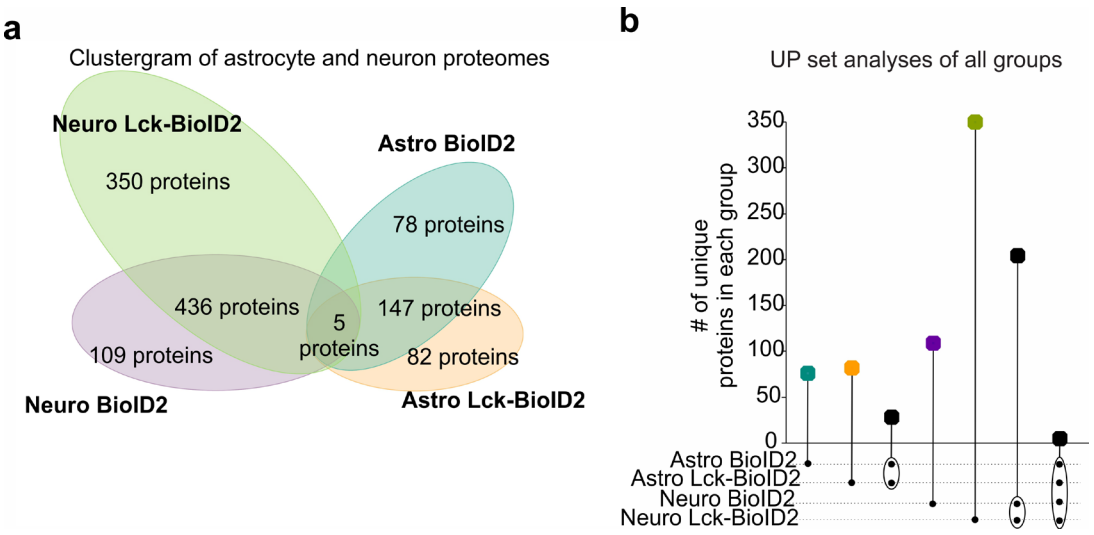

Figure 3.10: Broad assessment of cell and sub-compartment specific proteins. a. Clustergram shows number of unique proteins identified in each BioID2 construct experiment. The proteins represent those that were significant ($Log_2FC > 1$ and $FDR < 0.05$ *versus* GFP controls). **b.** UpSet plot of BioID2 identified proteins.

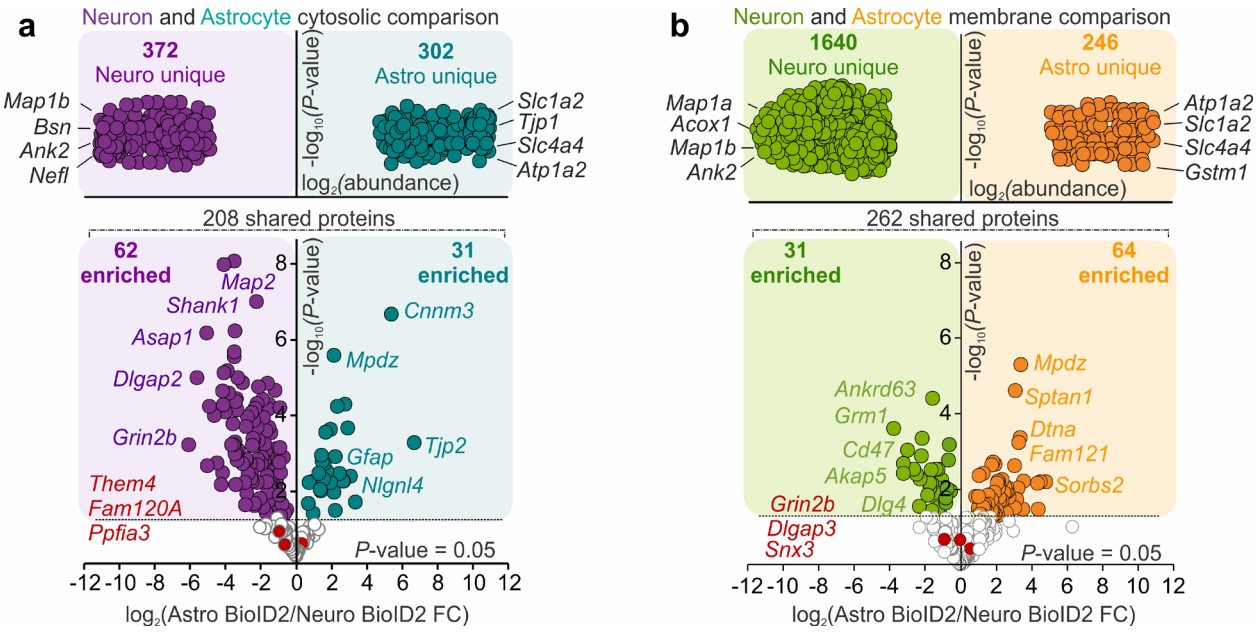

Figure 3.11: Quantitative comparison of astrocyte and neuron proteomes. a. Label-free based quantification comparison of proteins identified in both Astro BioID2 and Neuro BioID2 ($Log_2FC > 1$ and FDR < 0.05 *versus* GFP controls). Top plot shows unique proteins identified in Neuro BioID2 or Astro BioID2 when compared to each other. The four most abundant proteins are labeled. Lower volcano plot depicts the quantitative comparison of 208 proteins that were shared in both cytosolic Astro BioID2 and Neuro BioID2. The five highest enriched proteins in each cell-type are labeled. The top three proteins that showed no enrichment are labeled in red. **b.** Label-free based quantification comparison of proteins identified in both plasma membrane targeted Astro Lck-BioID2 and Neuro Lck-BioID2 ($Log_2FC > 1$ and FDR < 0.05 *versus* GFP controls). Top plot shows unique proteins identified in Neuro Lck-BioID2 or Astro Lck-BioID2 when compared to each other. The four most abundant proteins are labeled. Lower volcano plot depicts the quantitative comparison of 262 proteins that were shared in both plasma membrane Astro Lck-BioID2 and Neuro Lck-BioID2. The five highest enriched proteins in each cell-type are labeled. The top three proteins that showed no enrichment are labeled in red.

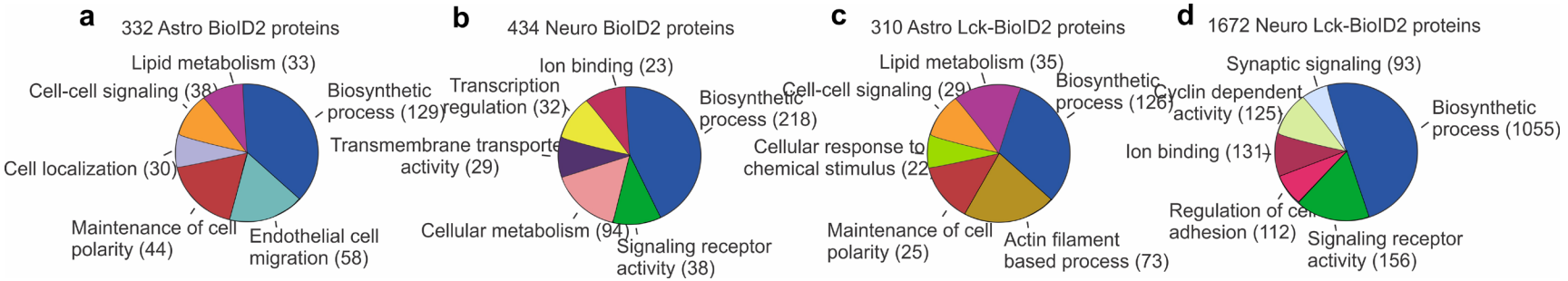

Figure 3.12: Broad functional analyses of neuron and astrocyte proteomes a. Pie chart of PANTHER pathway analysis terms for "biological processes". Pie chart shows the number of proteins found for each term from the 332 Astro BioID2 proteins. **b.** Pie chart of PANTHER pathway analysis terms for "biological processes". Pie chart shows the number of proteins found for each term from the 434 Neuro BioID2 proteins. **c.** Pie chart of PANTHER pathway analysis terms for "biological processes". Pie chart shows the number of proteins found for each term from the 310 Astro Lck-BioID2 proteins. **a.** Pie chart of PANTHER pathway analysis terms for "biological processes". Pie chart shows the number of proteins found for each term from the 1672 Neuro Lck-BioID2 proteins

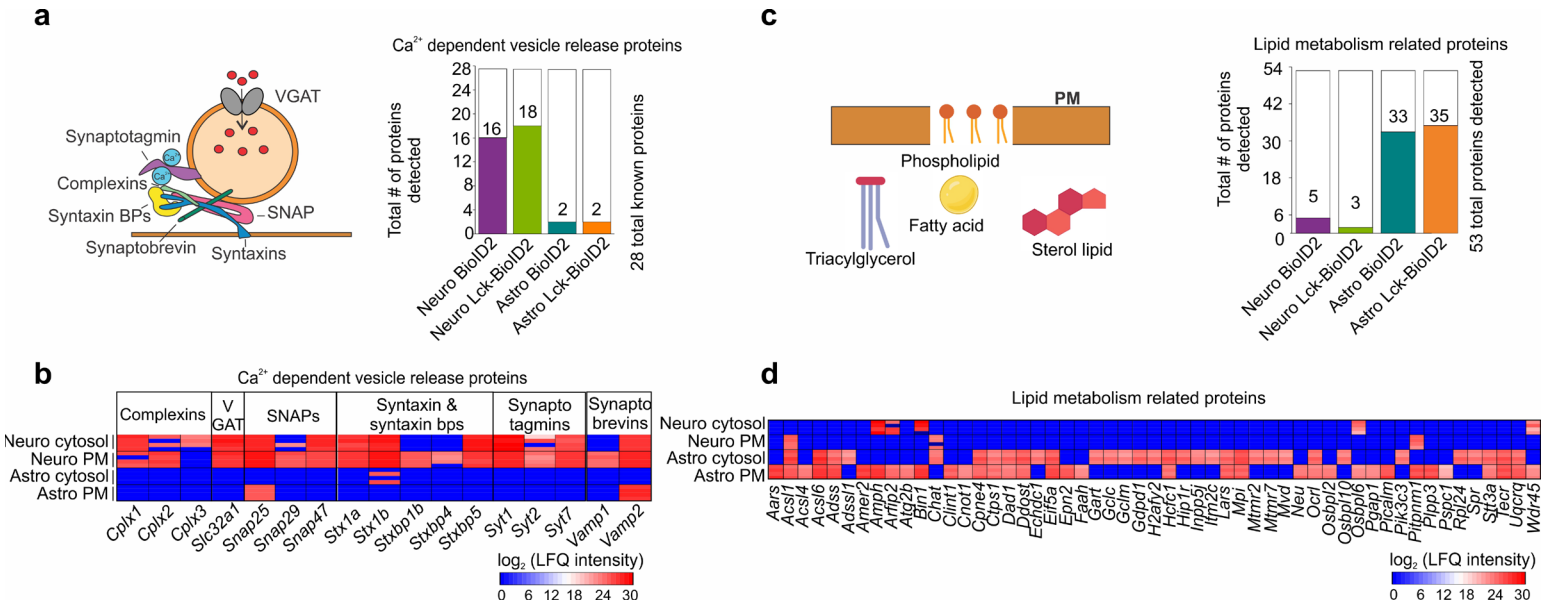

Figure 3.13: Assessment of specific neuron and astrocyte functions. a. Cartoon depicting the molecular components of Ca2+ dependent vesicular release required for neurotransmission. Bar graph denotes the number of Ca^{2+} dependent vesicle release protein isoforms that were detected in each BioID2 construct experiment. A total of 28 proteins are known. **b.** Expression levels (in label-free quantification intensity, LFQ intensity) of Ca^{2+} dependent vesicle release protein isoforms identified in each BioID2 construct (n = 4 experimental runs). Proteins in this analysis represent proteins that were enriched ($Log_2FC > 1$ versus GFP controls). Only proteins that were detected are shown. Scale shows Log_2LFQ intensity. **c.** Cartoon depicting the simplified process of lipid metabolism in astrocytes. Bar graph denotes the number of lipid metabolism related protein isoforms that were detected in each BioID2 construct experiment. A total of 53 proteins were detected in the BioID2 experiments. **d.** Expression levels of lipid metabolism related protein isoforms identified in each BioID2 construct (n = 4 experimental runs). Proteins in this analysis represent proteins that were enriched ($Log_2FC > 1$ versus GFP controls). Only proteins that were detected are shown.

datasets and *in vivo* proteomes. As mentioned before, the relationship between RNA transcripts and protein products is not straightforward, thus a comparison between proteomic and transcriptomic output could help us integrate and interpret these datasets more meaningfully. To make these datasets comparable, we used an AAV RiboTag method in the striatum to assess the transcriptomes of astrocytes and neurons for downstream comparison to our proteomic datasets.

Using a Rpl22-HA AAV construct which expresses a ribosomal subunit with an HA-tag in neurons or astrocytes selectively, we immunoprecipitated the tagged ribosomes and extracted the RNA from striatal neurons or astrocytes. We next conducted high-throughput deep sequencing to assess those transcripts we isolated. We compared the immunoprecipitated (IP) RNA to the RNA from the "soup of all cells" (also called the input sample) and by using a ratio of IP/input of greater than 1, we can detect astrocyte or neuron enriched transcripts. About 40% of the top 50 most enriched neuronal genes were also found at the protein level via our datasets. Furthermore, the top 50 neuronal genes were depleted in the astrocyte IP samples (**Figure 3.14**). Meanwhile, about 25% of the top 50 most enriched astrocyte genes were also found at the protein level in our datasets. These top 50 enriched genes were depleted in the neuronal IP samples (**Figure 3.14**). This underscores that a) neurons and astrocytes are fundamentally different and b) RNA does not accurately reflect protein product expression *in vivo.*

We next plotted the protein LFQ abundance from our datasets with the RNA fragments per kilobase per million (FPKM) abundance from the cell-specific RiboTag method. Across all proteomic datasets, the correlation was less than 0.35 (Pearson's *r*) and the slope of the line ranged from 0.26 to 0.38, indicating low correlation between RNA and protein abundances (**Figure 3.15**). Because sample preparation, data collection, and analysis methods differ between our transcriptomic and proteomic datasets, we then decided to conduct a correlation analysis based on

protein and RNA rank. The protein and RNA transcripts were ranked based on abundance and the protein rank was plotted as a function of the RNA rank. Similarly to the straightforward abundance correlations, this analysis showed that protein and RNA rank were very lowly correlated with a Pearson's *r* range from 0.26 to 0.38, thus indicating that with existing measurement methodology, RNA does not accurately reflect or predict protein product (**Figure 3.15**). Finally, using a Rank-rank Hypergeometric Overlap (RRHO) statistical analysis method (Plaisier et al., 2010) to show the strength and pattern of both RNA and protein expression profiles, we found that in both astrocyte and neuronal cytosol proteins, RNA can most accurately predict those gene products that are highly ranked in both RNA and protein levels. Interestingly, this is not the case with the plasma membrane proteomes in which there is no clear prediction of RNA to its corresponding protein (**Figure 3.15**). This may hint at some interesting biology in which plasma membrane proteins may have a different relationship with their corresponding transcripts such as varying turnover rates compared to other types of proteins in a cell.

Section 3.8: Case-by-case validation of proteins expressed by astrocytes

Given the richness of data presented here, we validated proteins that we found in our astrocyte expressing lists to assess astrocyte expression. While RNA does not correlate with protein, at least some RNA expression is needed to make protein product and this provides evidence of expression by a cell. Using *in situ* hybridization, we validated 3 proteins found within our astrocyte proteome datasets. We validated *Crym* mRNA expression, a striatum astrocyte marker; we validated *Mapt* which is canonically found in neurons, and we validated *Tjp1* which has been canonically associated with endothelial cells. First, using a single-cell RNA-seq data base from striatum (DropViz.org from Saunders et al., 2018) we quantified the number of predicted transcripts per

100,000 of each transcript in each cell type cluster including astrocytes, neurons, endothelial cells, and oligodendrocytes. We then conducted *in situ* hybridization of each gene in mouse striatum sections which contained tdTomato expressing astrocytes. We used immunohistochemical methods in tandem with the hybridization to both amplify the tdTomato signal and double label the astrocytes with S100β. We quantified the number of mRNA fluorescent puncta found within each tdTomato and S100 β positive astrocyte to assess expression. Our validation experiment found that striatal astrocytes do indeed express *Crym*, *Mapt*, and *Tjp1* at both the transcript and protein level (**Figure 3.16**). While *Mapt* and *Tjp1* have been canonically associated with neurons and endothelial cells, respectively, astrocytes do express these genes at both the transcript and protein level as shown with scRNA-seq, *in situ* hybridization, and proteomics. This highlights that perhaps associating expression of a gene product with a particular cell type to indicate cell identity might not be as straightforward as expected. While a protein can be enriched in a particular cell type, this gene can be expressed in another cell type where it may be conducting different molecular functions albeit at a lower expression level.

Section 3.8: The striatal astrocyte and neuron proteome networks

We combined both the astrocyte cytosol and the astrocyte plasma membrane proteome and the neuron cytosol and the neuron plasma membrane proteome and defined these as the high confidence astrocyte and neuron proteomes. To assess putative physical or functional interactions between the proteins identified, we used the STRING (Search Tool for the Retrieval of Interacting Genes/proteins) database which derives these functional predictions from five sources: genomic context predictions, high-throughput lab experiments, conserved co-expression, automated textmining, and previous knowledge in other databases. Through this analysis, we were able to

define the sources of the putative protein interactions so they could be limited by only genomic context predictions, lab experiments, and conserved co-expression. When we analyzed the 549 proteins that constituted the high-confidence astrocyte proteome with STRING, the PPI enrichment confidence value was 6.07×10^{-5} indicating that the interaction of these proteins with each other is not random and the observed number of interactions (edges) is significant. This indicates that these proteins are not only within the same spatial neighborhood, but that they may interact with one another to conduct diverse functions within astrocytes. We also analyzed the top 550 most abundant proteins that comprised the neuron proteome with STRING which gave a PPI enrichment confidence value of 7.07×10^{-5}. We created a map of both astrocyte and neuron proteomes (**Figure 3.17 and 3.18**) in which we could visualize the proteins and their putative interactions with one another. Using PANTHER gene ontology analysis, we then mapped the functional pathways that were most represented by the proteomes. The proteins comprising the astrocyte proteome mostly included proteins related to lipid metabolism, cell-cell signaling, and actin based processes (**Figure 3.17**). The neuronal proteome mostly consisted of proteins related to signaling receptor activity, synaptic signaling, and cellular metabolism (**Figure 3.18**). Taken together, our proteome maps reflect the diversity of proteins found in astrocytes and neurons, how they may interact, and how they may function. Thus, these results highlight that core signaling machinery is what may diverge astrocyte and neuron function in the striatum.

Section 3.9 Summary and Discussion

We have developed both cell and subcompartment proximity-dependent labeling tools to assess astrocyte and neuron proteomes *in vivo.* Tools to assess *in vivo* proteomes of intact cells are imperative to most accurately define and understand the function of these cells in defined central

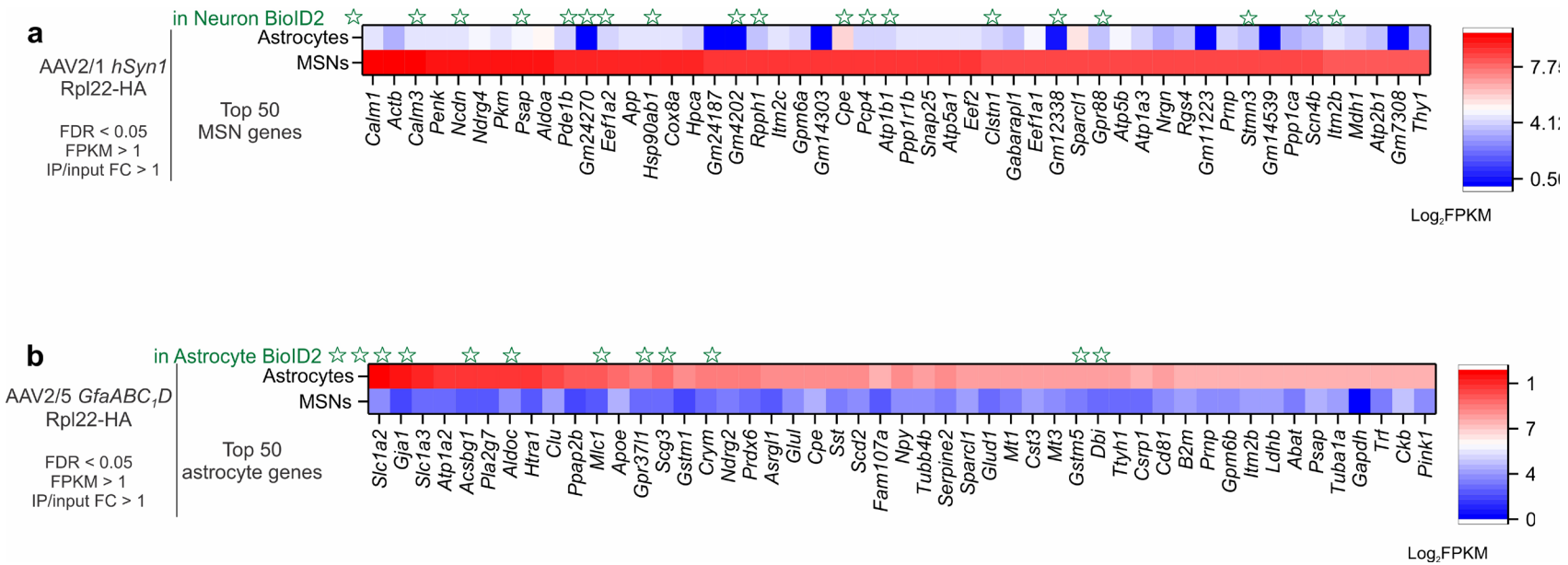

Figure 3.14: Transcriptomic analysis of striatal astrocytes and neurons. a. Top 50 medium spiny neuron (MSN) genes found by neuron-specific RiboTag AAV RNA-sequencing listed by the highest expression value (fragments per kilobase per million, FPKM). Astrocyte expression values for the top 50 MSN genes are shown. Green stars label gene products that were found at the proteomic level in Neuro BioID2 or Neuro Lck-BioID2. Scale shows Log_2FPKM. **b**. Top 50 astrocyte genes found by astrocyte-specific RiboTag AAV RNA-sequencing listed by the highest expression value (fragments per kilobase per million, FPKM). MSN expression values for the top 50 astrocyte genes are shown. Green stars label gene products that were found at the proteomic level in Astro BioID2 or Astro Lck-BioID2. Scale shows Log_2FPKM

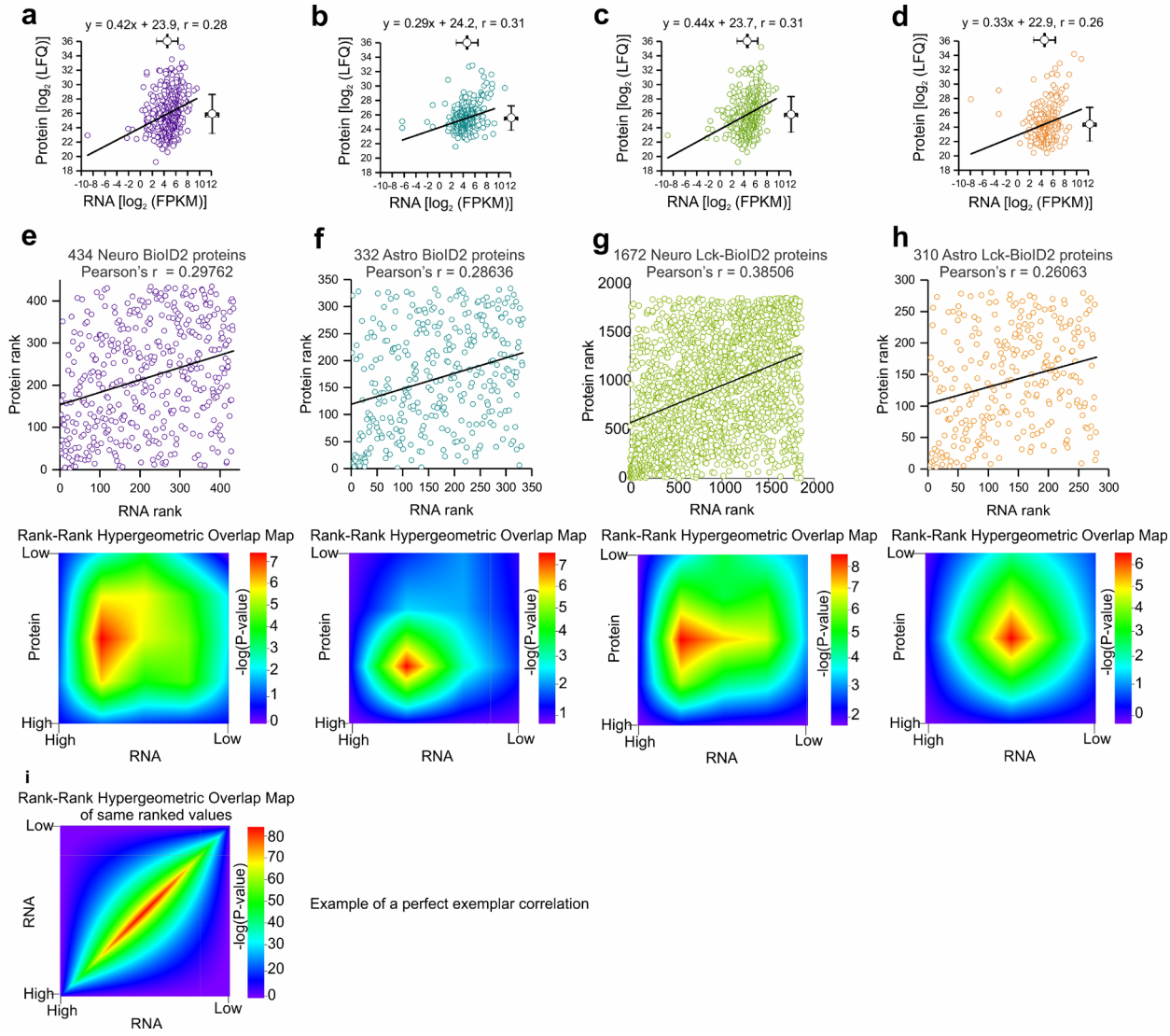

Figure 3.15: Transcriptome and proteome correlations. a. Scatter graph showing protein expression (Log_2LFQ intensity) as a function of mRNA expression (Log_2FPKM) for the 434 Neuro BioID2 proteins identified. Pearson's **r** is shown. **b.** As in **a**, but for the 332 Astro BioID2 proteins. **c.** As in **a**, but for the 1672 Neuro Lck-BioID2 proteins. **d.** As in **a**, but for the 310 Astro Lck-BioID2 proteins. **e**. Scatter graph showing the protein rank (by LFQ abundance) as a function of mRNA rank (by FPKM abundance) of the 434 Neuro BioID2 proteins identified. Pearson's **r** is shown. Heat map shows the rank-rank hypergeometric overlap (RRHO) of the RNA and protein rank. Each pixel represents the significance of overlap between the two datasets in $-\log_{10}$(P-value). Red pixels represent highly significant overlap. Color scale denotes the range of P-values at the negative $\log_{10}$ scale (Bin size = 100). **f.** As in **e**, but for Astro BioID2. **g.** As in **e**, but for Neuro Lck-BioID2. **h.** As in **e**, but for Astro Lck-BioID2. **i.** Example of a perfect correlation.

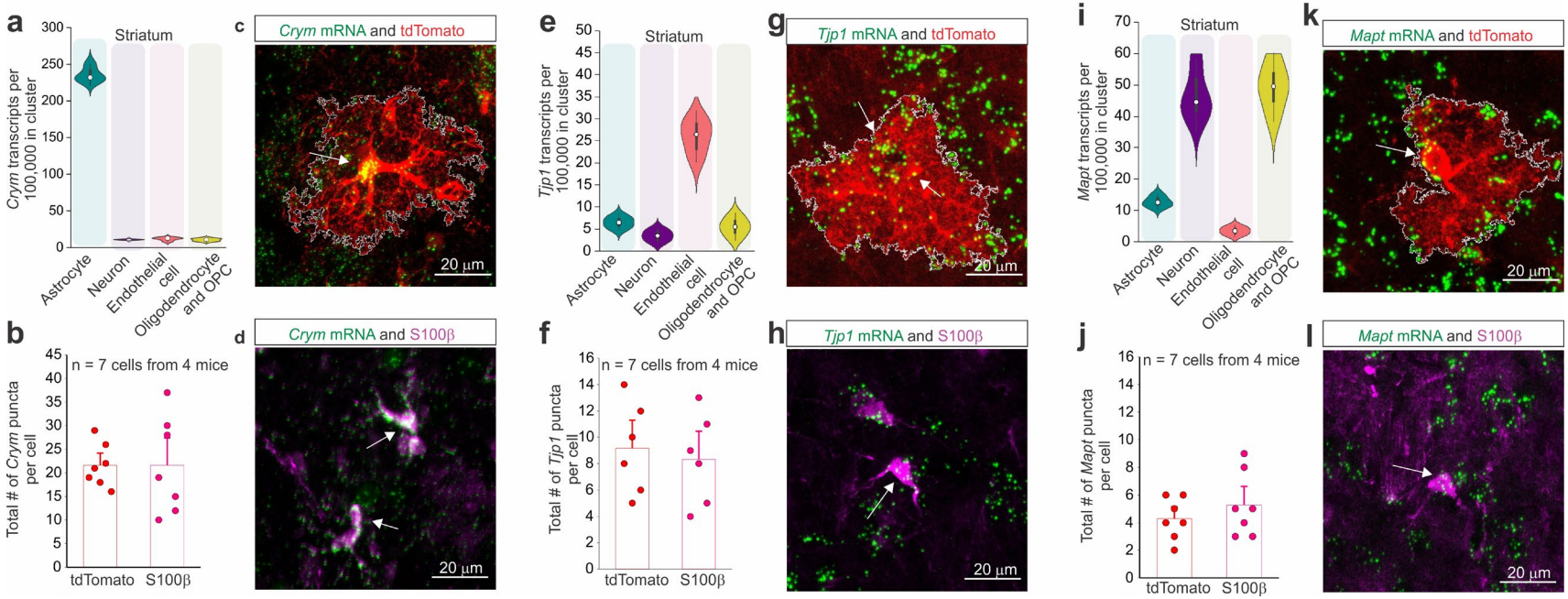

Figure 3.16: mRNA expression validation of astrocyte identified proteins. a. Violin plot from DropViz.org showing the expression of *Crym* in astrocytes, neurons, endothelial cells, and myelin associated cells of the striatum. **b.** Bar graphs denoting the total number of *Crym* mRNA puncta per tdTomato+ astrocyte and S100β+ cell. (n = 7 cells from 4 mice per group). **c.** Representative image of a tdTomato+ striatal astrocyte with *Crym* mRNA puncta expression (green). Arrow denotes puncta within the cell. **d.** Representative image of S100β+ cells with *Crym* mRNA puncta expression (green). Arrows denote the puncta within the cells. **e.** as in **a**, but for expression of *Tjp1*. **f.** as in **b**, but for *Tjp1* puncta. **g.** as in **c**, but for *Tjp1* mRNA puncta expression (green). **h.** as in **d**, but for *Tjp1* mRNA puncta expression (green). **i.** as in **a**, but for expression of *Mapt.* **j.** as in **b**, but for *Mapt* puncta. **k.** as in **c**, but for *Mapt* mRNA puncta expression (green). **l.** as in **d**, but for *Mapt* mRNA puncta expression (green).

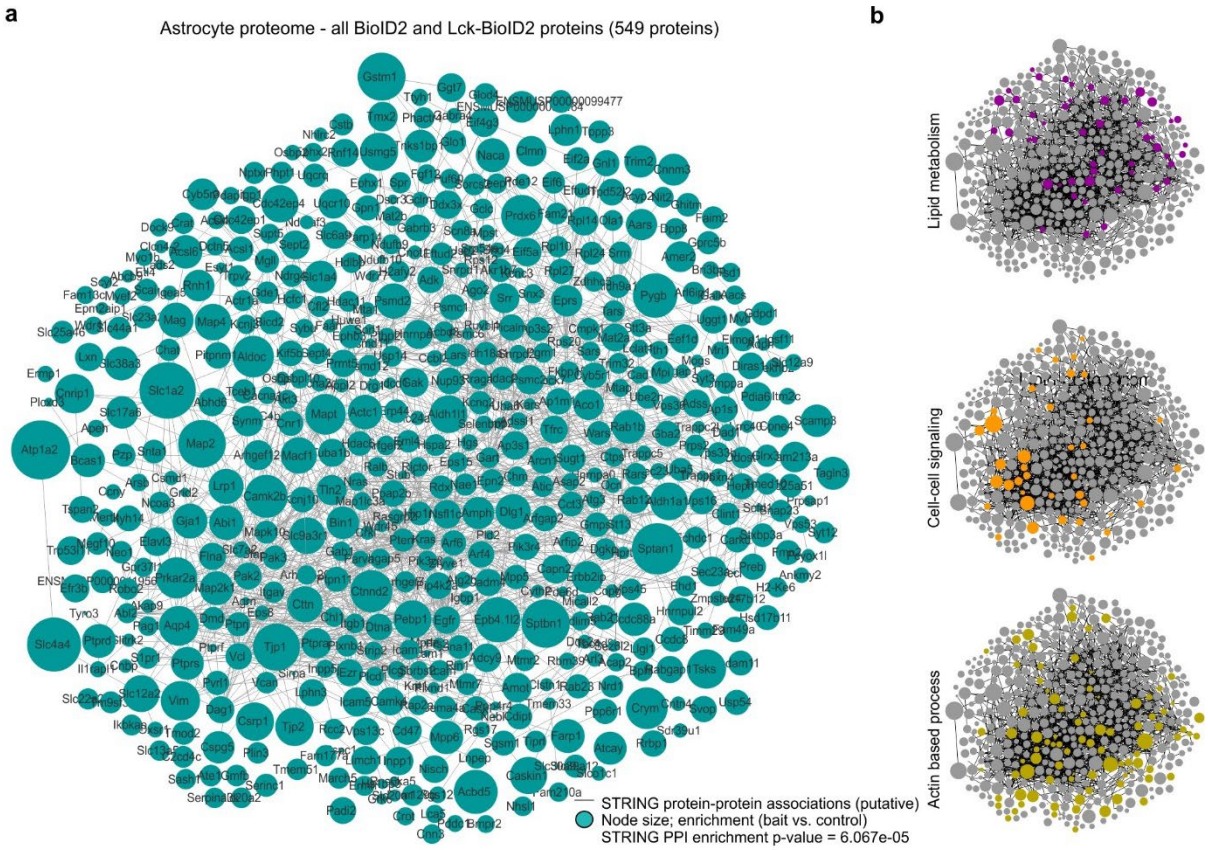

Figure 3.17: Interaction map for the astrocyte proteome. a. Scale-free STRING analysis interaction map of all the 549 unique and enriched biotinylated proteins identified in astrocytes with Astro BioID2 and Astro Lck-BioID2 . Node size represents the enrichment of each protein vs the GFP control (log2(BioID2/GFP)). Edges represent known interactions from the STRING database. **b**. Small clustergrams show selected functional categories from Panther GO analysis for "biological process" in different colors.

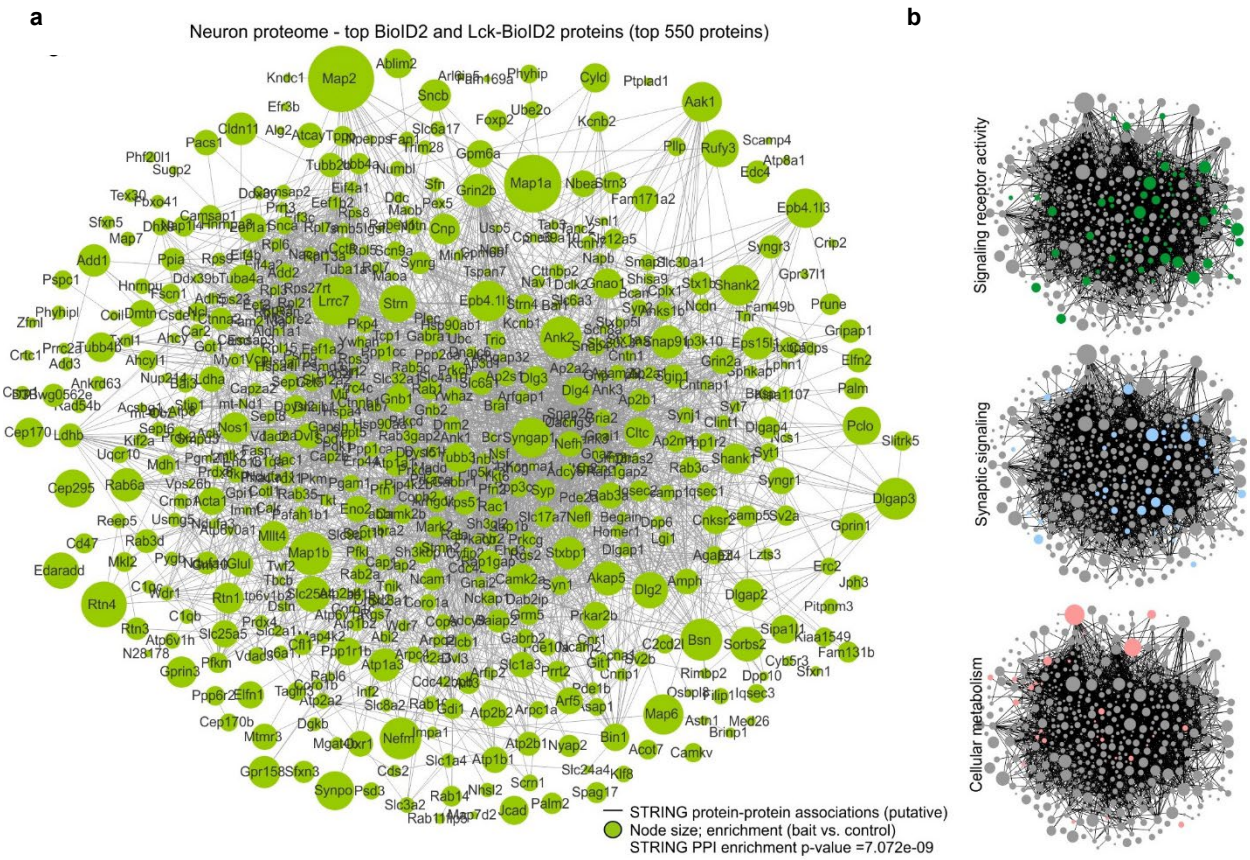

Figure 3.18: Interaction map for the neuronal proteome. a. Scale-free STRING analysis interaction map of the top 550 unique and enriched biotinylated proteins identified in neurons with Neuro BioID2 and Neuro Lck-BioID2 . Node size represents the enrichment of each protein vs the GFP control (log2(BioID2/GFP)). Edges represent known interactions from the STRING database. **b**. Small clustergrams show selected functional categories from Panther GO analysis for "biological process" in different colors.

nervous system circuitry. Here, we defined the high confidence proteomes of both astrocytes and neurons in the striatum where proteins found in neurons were most associated with synaptic signaling function and those in astrocytes conduct functions mostly related to lipid metabolism and homeostasis. We address three fundamental concepts with these studies.

First, we defined the astrocyte and neuron cytosolic and plasma membrane proteomes within the striatum for the first time. We learned that the core functions of neurons differ from the core functions of astrocytes by distinct cell-signaling pathways that defined their divergent signaling mechanisms. In neurons, cell-cell signaling is mostly defined by Ca^{2+} vesicle release whereas in astrocytes, core cell-cell signaling features related to their morphology by actin-signaling based processes and by homeostatic pathways such as lipid metabolism. Further analysis into these proteomes may identify cell-specific proteins that allow intercellular communication between astrocytes and neurons. Furthermore, hypothesis driven experiments with the use of these proteomes may allow for discovery of novel signaling receptors that could or may be involved in striatal dysfunction.

Second, we identified common proteins between both astrocyte and neuron proteomes which reveals interesting biology within the striatum. If there is a subset of proteins that are highly expressed in both astrocytes and neurons, this raises the question of whether these proteins have convergent or divergent functions in these two different cell types. Since we have established that indeed astrocytes and neurons are functionally different, what is then the relevance of these common proteins in both cell types? Interestingly, one of these proteins is a post-synaptic density associated protein, SAPAP3 (gene: *Dlgap3*), which we further investigate in latter chapters. Overall, the neuron and astrocyte proteomes reveal interesting biology which remains to be further explored.

Lastly, we show that for both astrocytes and neurons, mRNA transcript abundance cannot accurately predict protein abundance. By way of two different analyses including total abundance correlations and rank-rank correlations, we found that only the highest ranked proteins could be moderately predicted by RNA. This was an interesting result because the RNA we studied was RNA that was actively bound to ribosomes via the cell-specific RiboTag method and was most likely to be translated into protein. This indicates that post-translational regulation including modifications and protein turnover mechanisms are non-trivial and regulate a large aspect of a cell's biology. Thus, this is an important consideration to keep in mind when assessing transcriptomic data.

Overall, the neuron and astrocyte subcompartment proteomes we show here provide a molecular basis of function in both astrocytes and neurons in the striatum and also show that *in vivo* proteomic profiling in distinct subcompartments is possible using genetically targeted BioID2. These tools can be used in different physiological contexts and in distinct brain circuits

CHAPTER 4: ASTROCYTE SUBPROTEOMES

Section 4.1: Introduction

Astrocytes are morphologically complex cells with specialized finger-like processes that can contact other brain cells extensively. On average, astrocytes have complex bushy-like shapes with diameters of about 40-60 μm with volumes in the order of 10^4 μm^3 (Bushong et al., 2002; Ogata and Kosaka 2002; Chai et al., 2017; Salmon et al., 2023). Astrocytes interact with other cells at specialized contact points such as peripheral fine processes which interact with neuronal synapses, with blood vessels at polarized structures called end feet, or with other astrocytes at gap junctions. Therefore, their complex morphology allows astrocytes to conduct their various functions in the central nervous system at these diverse contact points. Investigating these specialized subcellular compartments within an astrocyte would thus, provide molecular mechanisms by which astrocytes conduct their diverse functions.

Given that BioID2 is small enough to be trafficked to small subcellular compartments (Kim et al., 2016) and that BioID2 can be targeted in a cell and subcompartment specific manner *in vivo* (**Chapter 3**), we developed a toolset to genetically target BioID2 to five specialized and functional astrocyte subcompartments. In this way, we could assess the protein-driven functions of striatal astrocyte at distinct contact points with other cells.

We chose five astrocyte subcompartments which represent core functions of astrocytes: 1) the polarized peri-vascular astrocyte end foot, 2) astrocyte gap junctions, 3) astrocyte fine processes, 4) astrocyte peri-glutamatergic synaptic processes, and 5) astrocyte processes at sites of potassium buffering. To genetically target BioID2 at these sites, we fused BioID2 to known proteins that are enriched and specially localized to these subcompartments. For the astrocyte peri-vascular endfoot we fused BioID2 to Aquaporin-4 (Nielsen et al., 1997; Mader and Brimberg, 2019) to create Aqp4-

BioID2. For astrocyte-astrocyte gap junctions, we fused BioID2 to Connexin-43 gap junctional channel (Spray, 1998; Giaume et al., 2010; Stout et al., 2015) to create Cx43-BioID2. We used actin associated protein, Ezrin, to assess astrocyte fine processes (Derouiche and Geiger, 2019) and create Ezr-BioID2. Assessment of the glutamate homeostasis subcompartment was done by using the astrocyte enriched glutamate transporter, Glt-1 (Lehre et al., 1998; O'Kane et al., 1999; Danbolt 2001) to create Glt-1-BioID2. Finally, to target the subcompartment related to potassium homeostasis, we used the astrocyte potassium channel, Kir4.1 (Bordey et al., 1999; Kofuji and Newman, 2004) to create Kir4.1-BioID2.

This palette of subcompartment specific BioID2 tools within astrocytes allowed us to assess and define the molecular basis by which astrocytes conduct their diverse physiological functions through their complex morphology. In this way, these subproteomes represent a bounty of novel information that can elucidate mechanisms driving pathophysiology during a variety of contexts including injury, stroke, or neurodegeneration. Furthermore, the astrocyte subproteomes also show expression of surprising proteins such as SAPAP3 (*Dlgap3)* in discrete functional compartments, allowing us to dissect its function in astrocytes via hypothesis testing. Thus, the astrocyte subproteomes are a novel dataset that can prove useful in the glial biology field.

Section 4.2: Validation of astrocyte subcompartment specific BioID2 constructs

While BioID2 can be localized to distinct subcellular compartments, it was imperative to assess whether our genetically targeted BioID2 toolset allowed for *in vivo* expression at the five defined subcompartments in astrocytes. We validated the expression and biotinylation activity of Aqp4-BioID2, Cx43-BioID2, Ezr-BioID2, Glt-1-BioID2, and Kir4.1-BioID2 in the striatum with immunohistochemistry and western blots. We microinjected each AAV construct unilaterally into

the dorsal striatum. After three weeks of expression, we subcutaneously injected biotin into the mice for seven days. Sixteen hours after the last injection, the mouse brain was processed for either immunohistochemistry or western blotting.

With immunohistochemistry we found that Aqp4-BioID2 was expressed mostly at the astrocyte end foot, as most of the BioID2-HA signal was peri-vascular. Additionally, when visualized with streptavidin, the biotinylated protein signal was also peri-vascular indicating that the proteins most proximal to Aqp4-BioID2 were indeed those localized to the end foot. To further validate that these immunohistochemical data were relevant to endogenous end feet, we immunolabeled non-injected brain striatal sections with anti-Aquaporin-4 antibody. Upon visualization, the endogenous Aqp-4 signal was reminiscent of both the Aqp4-BioID2 and the resultant biotinylated protein signal indicating that astrocyte end foot localized BioID2 is both expressed in its intended targeted subcompartment and that the proteins that are biotinylated are also those within the end foot subcompartment (**Figure 4.1**). To further assess the *in vivo* functionality of Aqp4-BioID2, we conducted western blot analyses for the biotinylated proteins using streptavidin conjugated horseradish peroxidase and chemiluminescence. These results showed that at the site of expression, Aqp4-BioID2 biotinylates a range of proteins across distinct molecular weights when compared to the non-expressing contralateral striatum (**Figure 4.1**). Thus, these biotinylated proteins represent the end foot subproteome of the astrocyte.

We conducted the same immunohistochemical and western blot analysis with fine process localized, Ezr-BioID2. In this case, the Ezr-BioID2 signal could be visualized as bushy and cloud-like reminiscent of astrocyte fine processes. Furthermore, the biotinylated signal arising from Ezr-BioID2 was also cloud-like and co-localized strongly with the Ezr-BioID2 signal. We validated these data by immunolabeling non-injected striatal sections with anti-Ezrin antibody to visualize

endogenous Ezrin staining within the striatum. Similar to both the Ezr-BioID2 and biotinylated signal, the endogenous Ezrin signal was also cloud-like and bushy indicating that Ezr-BioID2 is localized to astrocyte fine processes (**Figure 4.2**). We assessed Ezr-BioID2 activity in the striatum by a western blot of biotinylated proteins. By these western blots, we found that Ezr-BioID2 indeed biotinylates proteins of different molecular weights and is therefore, is able to biotinylate proteins at astrocyte fine processes (**Figure 4.2**).

We repeated this immunochemical and western blot strategy with Glt-1-BioID2 which is localized to astrocyte processes at sites where astrocytes contact glutamatergic synapses. Both the Glt-1-BioID2 and its subsequent biotinylated protein signals were found to be punctate, but bushy and expressed at the astrocyte plasma membrane (**Figure 4.3**). When we assessed the endogenous Glt-1 localization pattern via anti-Glt-1 antibody labeling in separate striatal sections, we also saw punctate and bushy structures throughout the striatum similar to Glt-1-BioID2. Via western blot of the biotinylated proteins, Glt-1-BioID2 biotinylated a large range of proteins compared to the contralateral control (**Figure 4.3**). Thus, Glt-1-BioID2 is indeed expressed at sites where astrocytes contact glutamatergic synapses and functionally biotinylates proteins at these sites.

We then validated Kir4.1-BioID2, which localizes to astrocytic sites of potassium buffering, a major homeostatic role that astrocytes play in the central nervous system. Using immunohistochemistry we found that Kir4.1-BioID2 and the biotinylated protein signals had a similar pattern of expression as Aqp4-BioID2. This was unsurprising as Kir4.1 and Aqp4 endogenously associate with one another (Nagelhus et al., 2004). Additionally, we were able to visualize more bushy structures that had a similar semblance to astrocyte processes thus indicating that Kir4.1-BioID2 was also expressed at astrocyte processes where potassium buffering occurs (**Figure 4.4**). We assessed endogenous Kir4.1 localization with anti-Kir4.1 antibody labeling and

found that similar to Kir4.1-BioID2, endogenous Kir4.1 is localized to astrocyte processes and to end feet. Western blot of striatum expressing Kir4.1-BioID2 demonstrated that Kir4.1-BioID2 also biotinylates proteins *in vivo* when compared to the contralateral side indicating that the potassium buffering subproteome could be subsequently characterized (**Figure 4.4**).

Finally, we validated astrocyte gap junction localized Cx43-BioID2 with both immunohistochemistry and western blot. Gap junctions are characterized by their formation of plaques which can be visualized as discrete puncta. Immunostaining of Cx43-BioID2 showed plaque-like puncta indicating that Cx43-BioID2 is localized to astrocyte gap junctions with other astrocytes. Interestingly, the biotinylated protein pattern showed both these distinct puncta, and proteins in the astrocyte somata (**Figure 4.5**). This indicated that proteins that interact with astrocyte gap junctions may move to other locations to conduct functions in other astrocyte subcompartments. Validation of endogenous Cx43 expression patterns with an anti-Cx43 antibody in the striatum showed similar plaque-like puncta as the Cx43-BioID2 pattern, indicating that Cx43-BioID2 is indeed localized to astrocyte gap junctions. Western blot analysis of Cx43-BioID2 biotinylated proteins showed that Cx43-BioID2 is functional as a large range of proteins was biotinylated when compared to the non-expressing contralateral control. Therefore, Cx43-BioID2 is localized and functional at sites of astrocyte-astrocyte contact (**Figure 4.5**).

Section 4.3: Mass spectrometry based identification and analysis of astrocyte subproteomes

We have demonstrated that our palette of astrocyte and subcompartment tools for assessing proteomes are functional *in vivo* and are specific to distinct subcellular compartments. Therefore, we next used mass spectrometry to identify and understand the proteins found at these astrocyte functional subcompartments. We injected a set of mice with either sub-compartment specific

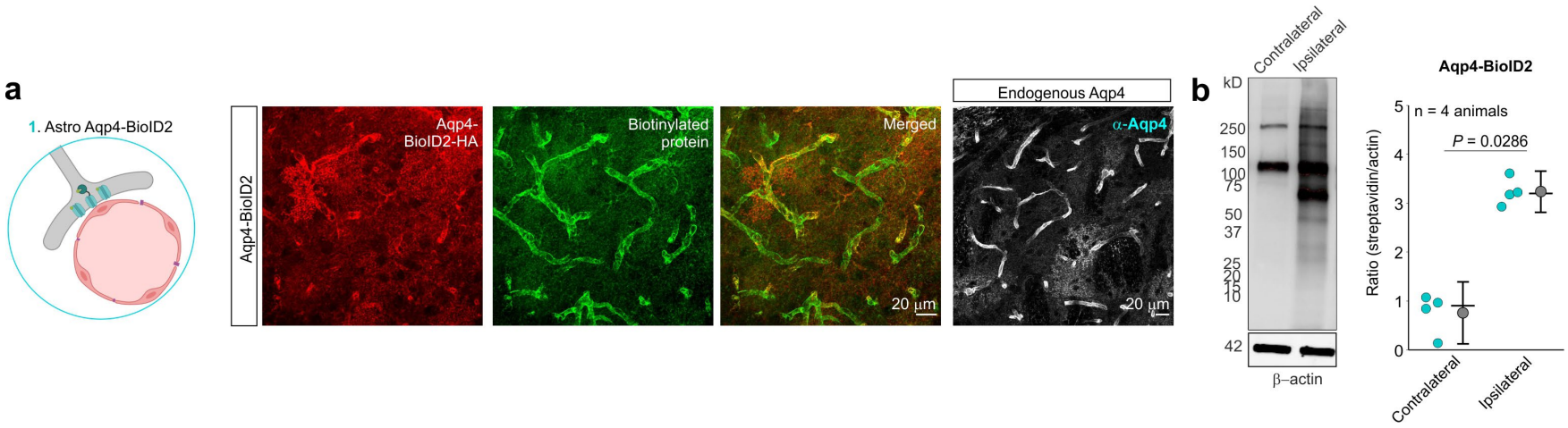

Figure 4.1: Validation of Aqp4-BioID2. a. Representative images of immunostained mouse striatum injected with astrocyte specific Aquaporin4-BioID2 (Aqp4-BioID2) and then treated with biotin for 7 days. The tissue was immunostained with both anti-HA antibody (red) and a fluorophore conjugated streptavidin probe (green). Farthest panel on the right shows the endogenous Aqp4 expression pattern in the striatum after immunostaining with anti-Aqp4 antibody. **b.** Western blot analysis of brain unilaterally microinjected with astrocyte specific Aqp4-BioID2. Dark bands at 130 kD and 250 kD show the endogenously biotinylated mitochondrial proteins, Pyruvate carboxylase and acetyl-CoA carboxylase. Graph depicting the streptavidin signal intensity divided by the β-actin signal intensity for each data point. Black horizontal line depicts the mean. (n = 4 mice; Mann-Whitney test)

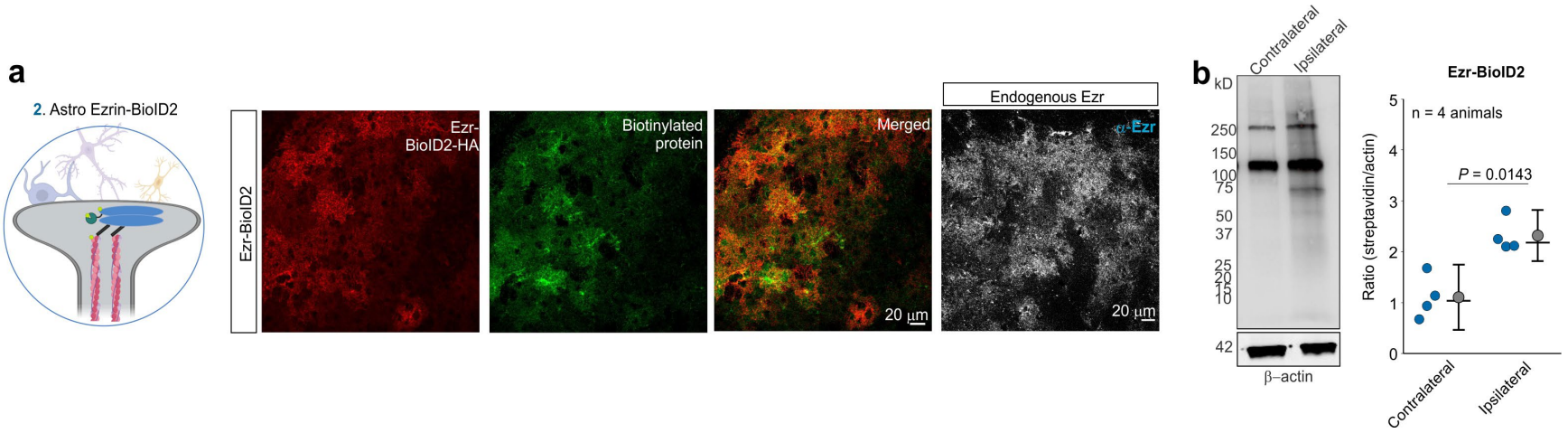

Figure 4.2: Validation of Ezr-BioID2. a. Representative images of immunostained mouse striatum injected with astrocyte specific Ezrin-BioID2 (Ezr-BioID2) and then treated with biotin for 7 days. The tissue was immunostained with both anti-HA antibody (red) and a fluorophore conjugated streptavidin probe (green). Farthest panel on the right shows the endogenous Ezr expression pattern in the striatum after immunostaining with anti-Ezr antibody. **b.** Western blot analysis of brain unilaterally microinjected with astrocyte specific Ezr-BioID2. Graph depicting the streptavidin signal intensity divided by the β-actin signal intensity for each data point. Black horizontal line depicts the mean. (n = 4 mice; Mann-Whitney test).

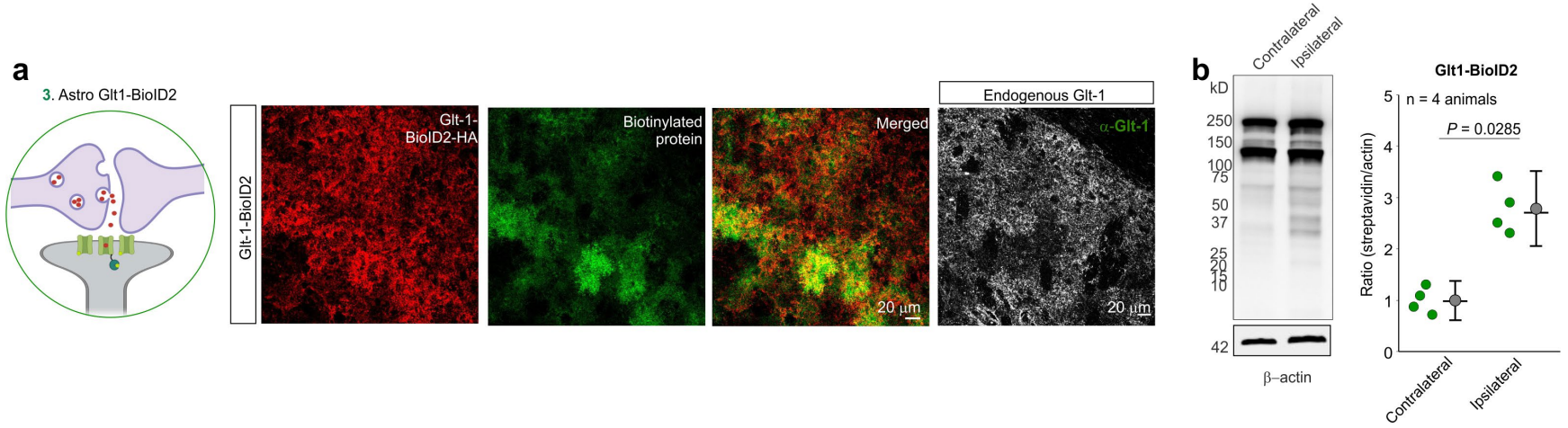

Figure 4.3: Validation of Glt1-BioID2. a. Representative images of immunostained mouse striatum injected with astrocyte specific Glt1-BioID2 and then treated with biotin for 7 days. The tissue was immunostained with both anti-HA antibody (red) and a fluorophore conjugated streptavidin probe (green). Farthest panel on the right shows the endogenous Glt-1 expression pattern in the striatum after immunostaining with anti-Glt-1 antibody. **b.** Western blot analysis of brain unilaterally microinjected with astrocyte specific Glt1-BioID2. Graph depicting the streptavidin signal intensity divided by the β-actin signal intensity for each data point. Black horizontal line depicts the mean. (n = 4 mice; Mann-Whitney test).

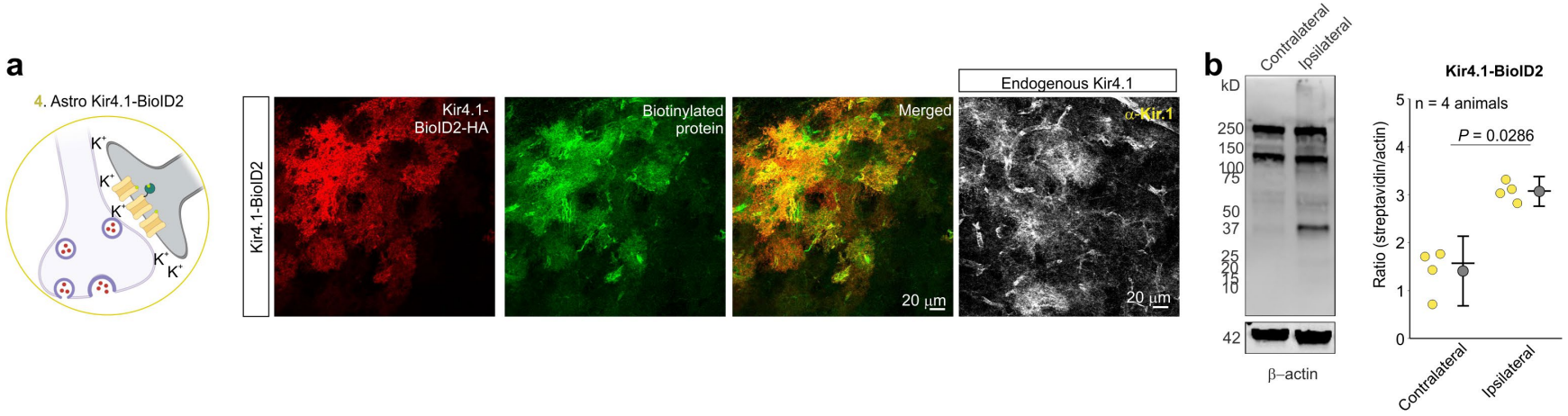

Figure 4.4: Validation of Kir4.1-BioID2. a. Representative images of immunostained mouse striatum injected with astrocyte specific Kir4.1-BioID2 and then treated with biotin for 7 days. The tissue was immunostained with both anti-HA antibody (red) and a fluorophore conjugated streptavidin probe (green). Farthest panel on the right shows the endogenous Kir4.1 expression pattern in the striatum after immunostaining with anti-Kir4.1 antibody. **b.** Western blot analysis of brain unilaterally microinjected with astrocyte specific Kir4.1-BioID2. Graph depicting the streptavidin signal intensity divided by the β-actin signal intensity for each data point. Black horizontal line depicts the mean. (n = 4 mice; Mann-Whitney test).

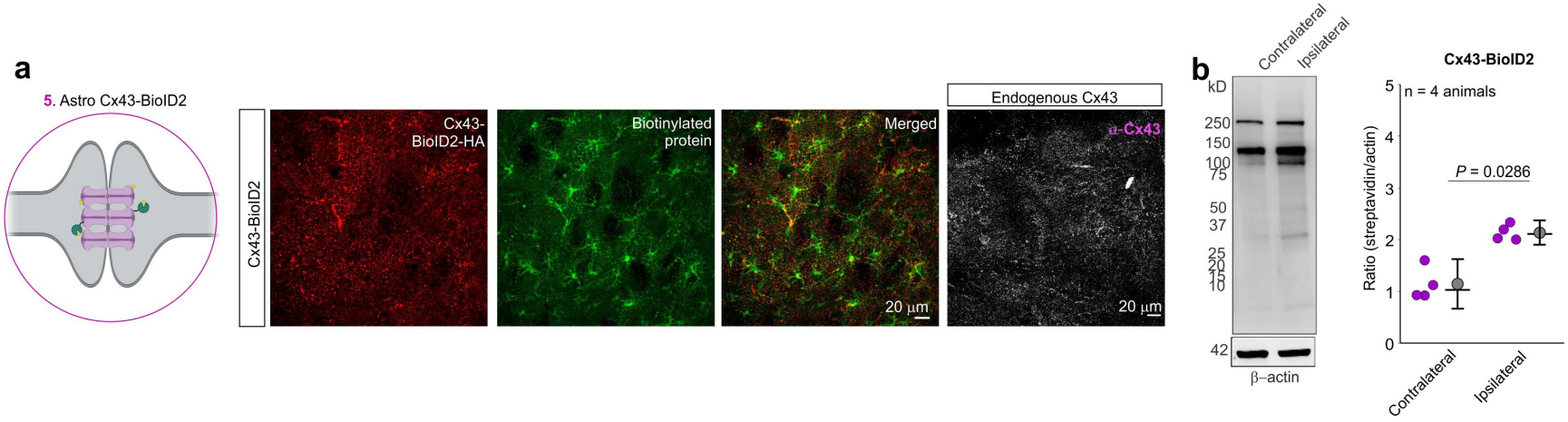

Figure 4.4: Validation of Cx43-BioID2. a. Representative images of immunostained mouse striatum injected with astrocyte specific Connexin43-BioID2 (Cx43-BioID2) and then treated with biotin for 7 days. The tissue was immunostained with both anti-HA antibody (red) and a fluorophore conjugated streptavidin probe (green). Farthest panel on the right shows the endogenous Cx43 expression pattern in the striatum after immunostaining with anti-Cx43 antibody. **b.** Western blot analysis of brain unilaterally microinjected with astrocyte specific Cx43-BioID2. Graph depicting the streptavidin signal intensity divided by the β-actin signal intensity for each data point. Black horizontal line depicts the mean. (n = 4 mice; Mann-Whitney test).

BioID2 AAVs bilaterally into the striatum. We also injected another set of mice with a cognate subcompartment-specific GFP AAV control such as Aqp4-GFP, Ezr-GFP, Glt1-GFP, Kir4.1-GFP, or Cx43-GFP as a control. After three weeks of expression, all mice received the same subcutaneous injection of 24 mg/kg of biotin for 7 days, after which the mouse striata were processed for purification of biotinylated proteins.

After purification of the proteins with streptavidin conjugated beads, the proteins were trypsinized to yield peptide fragments for downstream mass spectrometry analyses. We used high performance liquid chromatography coupled with tandem mass spectrometry to assess the identity and relative abundance of these cell specific peptide fragments. A total of 4 technical and biological runs consisting of 8 mouse striata each were run for each astrocyte subcompartment experiment in order to generate high confidence proteins lists.

We identified 4018 proteins in the combined Aqp-BioID2 and Aqp4-GFP mass spectrometry experiment, 3040 proteins in the combined Ezr-BioID2 and Ezr-GFP experiment, 3522 proteins in the Glt1-BioID2 and Glt1-GFP combined experiment, 3678 proteins in the Kir4.1-BioID2 and Kir4.1-GFP experiment, and 3377 proteins in the combined Cx43-BioID2 and Cx43-GFP experiment. The proteins identified here represent proteins that were identified in at least one technical run. We normalized the LFQ intensity values among technical runs by using the LFQ values of the endogenously biotinylated 130 kDa mitochondrial protein, Pyruvate carboxylase, which appeared robustly at similar abundance levels in all samples, including in the fluorescent protein controls.

After normalization with Pyruvate carboxylase, we conducted a dimensionality reducing analysis using principal component analysis (PCA) of the average intensity of all the proteins identified across all 4 technical replicates in order to visualize how each subcompartment-specific

BioID2 and their corresponding fluorescent protein control experiment related to each other. From this analysis, we found that principal component 1 explained about 28% of the variation of each construct sample and that the fluorescent protein controls separated far on this plane from all BioID2 experiments (**Figure 4.6**). We then manually removed all mitochondrial associated proteins as these can be transiently and endogenously biotinylated. Additionally, we removed proteins labeled by Uniprot as known contaminants such as mouse keratins. This left us with 3209 proteins in the combined Aqp4-BioID2 and Aqp4-GFP identification list, 2268 proteins in the combined Ezr-BioID2 and Ezr-GFP identification list, 2936 proteins in the Glt1-BioID2 and Glt1-GFP identification list, 3099 proteins in the Kir4.1-BioID2 and Kir4.1-GFP list, and 2914 proteins in the Cx43-BioID2 and Cx43-GFP list. Using the LFQ intensity and R statistical package, limma, we chose high confidence proteins that had larger than 2-fold LFQ intensity over the fluorescent protein controls, had a Bonferroni false discovery rate (adjusted p-value) of less than 0.05, and were present in three or more technical runs. This analysis yielded 806 proteins in the Aqp4-BioID2 identification list, 403 proteins in the Ezr-BioID2 identification list, 757 proteins in the Glt1-BioID2 identification list, 667 proteins in the Kir4.1-BioID2 identification list, and 295 proteins in the Cx43-BioID2 identification list. These values and filters are found in **Table 4.1** and **supplementary excel file 2**.

Section 4.4: Characterization of astrocyte subproteomes

Mass spectrometry and resulting statistical analyses allowed us to generate a high confidence list of proteins for each astrocyte subcompartment specific BioID2 experiment. These high confidence lists represent proteins that were biotinylated by BioID2 in discrete astrocyte subcompartments *in vivo*. To assess astrocyte-selectivity of subcompartment specific BioID2, we

asked whether proteins that are considered cell-enriched proteins for multiple brain cell types were present or absent from each astrocyte subcompartment specific experiment. For neurons we chose NeuN (gene: *Rbfox3*), Nefl (gene: *Nefl*), β-III tubulin (gene: *Tubb3*), and Enolase-2 (gene: *Eno2*). For astrocytes we chose Aldh1*l*1 (gene: *Aldh1l1*), Kir4.1 (gene: *Kcnj10*), Glt-1 (gene: *Slc1a2*) and Cx-43 (gene: *Gja1*). For oligodendrocytes we chose MOG (gene: *Mog)*, Sox10 (gene: *Sox10*), MOBP (gene: *Mobp*), and MBP (gene: *Mbp*). For oligodendrocyte progenitor cells we chose Olig2 (gene: *Olig2*), PDGFR-α, (gene: *Pdgfra*) and Cacng4 (gene: *Cacng4*). For microglia we chose CD-68 (gene: *Cd68*), Ptprc (gene: *Ptprc*), CCL-3 (gene: *Ccl3*), and TNF (gene: *Tnf*). When we searched our high confidence subcompartment protein lists for these cell-enriched markers we found that none of the astrocyte subcompartments we assessed showed any presence of cell enriched markers for neurons, oligodendrocytes, nor microglia, indicating that astrocyte and subcompartment specific BioID2 is indeed cell-selective. Interestingly, not all cell-enriched marker proteins for astrocytes were found in all astrocyte subcompartments. For example, Aldh1*l*1 was found in all subcompartments except in the Cx43 gap junctional compartment. Furthermore, Cx43 was not found in the Kir4.1 or Glt-1 subcompartments. These results indicate that although BioID2 was expressed in astrocytes, each genetically targeted construct was indeed localized to distinct subcompartments within the astrocytes where they biotinylated specialized subcompartment proteins (**Figure 4.7**). Taken together, these analyses support the immunohistochemical data that our astrocyte subcompartment specific BioID2 AAV constructs are both astrocyte-selective and sub-compartment targeted.

Comparison of our high confidence protein lists to each other revealed common proteins and proteins unique to each astrocyte subcompartment. There were 26 proteins shared by all subcompartments including the astrocyte plasma membrane and astrocyte cytosol. 247 proteins

were unique to the Aqp4 endfoot, 51 were unique to the Cx43 gap junction, 79 were unique to the Ezrin fine processes, 205 were unique to the Glt1 sites of glutamate uptake, 119 were unique to the Kir4.1 sites of potassium buffering, 115 were unique to the astrocyte plasma membrane, and 59 were unique to the cytosol (**Figure 4.8**). To generate a high confidence subproteome protein list for each subcompartment, we next compared the high confidence proteins for each subcompartment to the proteins found at the astrocyte cytosol with Astro BioID2. Using the LFQ intensity, and R statistical analysis, limma, we quantitatively compared the high confidence proteins of the cytosol to each subcompartment.

When we compared the astrocyte plasma membrane identified proteins to the cytosol, we found 238 proteins that were unique to the plasma membrane. Additionally, there were 144 proteins that were shared among the plasma membrane and the cytosol. Of these, 32 were significantly enriched in the plasma membrane. Among these was an electrogenic sodium bicarbonate co-transporter (*Slc4a4)* and surprisingly, post-synaptic density associated protein, SAPAP3 (gene: *Dlgap3*). Therefore, the 238 plasma membrane unique proteins in addition to the 32 plasma membrane enriched proteins constitute the high confidence astrocyte plasma membrane proteome (**Figure 4.9**).

When we compared the astrocyte Ezrin subcompartment with the cytosol, there were 216 proteins that were unique to Ezrin including known interactors of Ezrin such as Moesin (gene: *Msn*). 186 proteins were shared between Ezrin and the cytosol, of which 35 were significantly enriched in the Ezrin subcompartment. Among these was Radixin (gene: *Rdx*), which is also a known interactor of Ezrin and part of the well characterized ezrin-radixin-moesin (ERM) complex (Fohen et al., 2010) and again, SAPAP3 (gene: *Dlgap3*). Together, the 216 unique Ezrin proteins

in conjunction with the 35 enriched proteins make up the high confidence astrocyte Ezrin fine process proteome (**Figure 4.9**).

We next compared the Aqp4 endfoot subcompartment proteins with the cytosol. This comparison yielded 577 proteins that were unique to the Aqp4 endfoot. These included heat shock protein 4 (gene: *Hspa4*) which regulates protein localization. Of the 228 proteins shared between the Aqp4 endfoot and the cytosol, 90 were enriched in the end foot including proteins related to carbohydrate metabolism such as mannose-6-phosphate isomerase (gene: *Mpi*). Thus, the sum of the 577 unique Aqp4 endfoot proteins with the 90 enriched proteins represent the high confidence astrocyte Aqp4 endfoot proteome (**Figure 4.9**).

When we compared the Glt-1 glutamate transporter subcompartment proteins with the cytosol, we found 527 unique proteins in the Glt-1 compartment including actin binding molecules such as Cdc42 (gene: *Cdc42)*. There were 230 shared proteins between the Glt-1 compartment and the cytosol, of which 5 were enriched including Glt-1 (gene: *Slc1a2*) itself and a proton coupled zinc transporter (gene: *Slc30a1*). Together, the 527 unique proteins with the 5 enriched proteins make up the high confidence astrocyte Glt-1 glutamate transport proteome (**Figure 4.9)**

Comparison of the Kir4.1 potassium buffering subcompartment proteins with the cytosol revealed 390 Kir4.1 unique proteins. These included astrocyte adhesion molecule Glialcam (gene: *Hepacam*). 275 proteins were shared between Kir4.1 and the cytosol, of which only 3 were enriched in the Kir4.1 potassium buffering compartment including Kir4.1 itself and Aqp4. Thus, the 275 unique proteins in conjunction with the 3 enriched proteins are the high confidence Kir4.1 potassium buffering proteome in astrocytes (**Figure 4.9**).

Finally, we compared astrocyte gap junctions via Cx43 with the cytosol. In this analysis there were 179 Cx43 gap junction unique proteins including actin related protein, Arpc1a (gene:

Arpc1a). Of the 115 shared proteins between the cytosol and astrocyte gap junctions, 15 were enriched in the Cx43 gap junctions including Cx43 itself and tight junctional proteins ZO-1 (gene: *Tjp1*) and ZO-2 (gene: *Tjp2*), which have been described as important scaffolding proteins for cell adhesion. Thus, the 115 unique proteins summed with the 15 enriched proteins constitute the high confidence astrocyte Cx43 gap junction proteome (**Figure 4.9**).

In order to broadly assess the functions of the proteins found in each astrocyte subcompartment, we employed the Enrichr gene set analysis tool to assign broad functions to each subproteome. We took the most significant term for "biological process" and found that the plasma membrane subproteome consisted of proteins involved in "transport across the blood-brain barrier." The astrocyte end foot subproteome consisted of proteins involved in "vascular transport," the astrocyte Ezr fine processes were made up of proteins related to "plasma membrane bounded cell projection," the astrocyte Glt-1 subproteome included proteins related to "regulation of glutamatergic synaptic transmission," the Kir4.1 subproteome consisted of proteins related to "protein transfer" and the astrocyte gap junctional proteins were related to the "regulation of exocytosis" (**Figure 4.10**). When we conducted the same analysis, but for gene ontology related to "molecular function" we found that the astrocyte plasma membrane subproteome consisted of proteins related to "cadherin binding." The astrocyte end foot subproteome was made up of proteins relevant to "nucleoside tri-phosphatase activity," the astrocyte Ezr fine process subproteome was also related to "cadherin binding," the Glt-1 subproteome was involved in "purine ribonucleoside triphosphate binding," the Kir4.1 subproteome contained proteins also related to "nucleoside tri-phosphatase activity," and finally, the Cx43 gap junction subproteome was composed of proteins related to "protein serine kinase activity" (**Figure 4.10**). These gene ontology lists provide a couple of broad clues related to astrocyte subcompartment function.

First, is that subsets of subproteomes may have similar functions relevant to their localization. For example, both the plasma membrane subproteome and the Ezrin fine process subproteome contain proteins related to cadherin binding. Cadherins are a superfamily of cell-cell adhesion molecules that are necessary for maintenance of cell architecture and cell-cell contacts. Though cadherins contain both intracellular and extracellular domains, binding through cadherin extracellular domains is relatively weak, but strong cell-cell adhesion occurs by lateral clustering of cadherins by proteins that link the cadherin cytoplasmic domain to the actin cytoskeleton (Perez and Nelson 2004). One of the proteins involved in actin scaffolding is in fact, Ezrin. Therefore, it is unsurprising that both the astrocyte plasma membrane and the Ezrin fine processes both contain proteins that are relevant for cell-cell contacts, such as cadherins, and are also involved in anchoring or binding of such transmembrane proteins. Another example of such common functions occurs with the Aqp4 end foot subproteome and the Kir4.1 potassium buffering subproteome. As stated before, Aqp4 and Kir4.1 proteins have been shown to spatially interact (Nagelhus et al., 2004). In this context, Aqp4 protein was found within the Kir4.1 domain, and Kir4.1 was found within the Aqp4 end foot domain. Therefore, it was also unsurprising that both subproteomes contained proteins involved in similar molecular functions such as second messenger "nucleoside-triphosphatase activity."

Second, although some astrocyte subproteomes share similar functions, the functions defined by "biological processes" are diverse and thus indicate both the subcompartment specificity of genetically targeted BioID2 and the diversity of astrocyte subcompartment functions. Even within a complex astrocyte, proteins are specifically localized to domains where specialized functions are conducted and contacts with a variety of cell types are made. The most divergent biological

processes included the end foot and the Glt-1 subproteome which were involved in “vascular transport” and regulation of “glutamatergic synapse transmission”, respectively.

In all, the astrocyte subproteomes present a bounty of novel information that provides clues on what astrocytes in the striatum do and how they do it. Case-by-case studies of subsets of these proteins could provide a platform to conduct robust hypothesis driven experiments.

Section 4.5: Astrocyte subproteome cards

The astrocyte subproteome data presented here provides a plethora of unprecedented data that provides insight on astrocyte function in the striatum. While the data is rich in detail, it was imperative for us to present the data in a succinct, summarized, and easy to follow manner. To do this, we have created the “astrocyte subproteome cards” to summarize key details related to each astrocyte subcompartment in a standard manner. Each subproteome card contains six key pieces of information relevant to each subcompartment. First, (a) we use a schematic of each astrocyte subcompartment to detail which subproteome is the focus of the card, (b) we then show the volcano plot from **figure 4.9** of the comparison between the single BioID2-targeted subcompartment and cytosolic BioID2 to find subcompartment enriched proteins. Third, (c) we show the relationship between the protein abundance of the proteins found within that subcompartment to the RNA abundance from RiboTagged striatal astrocytes using a rank-rank hypergeometric overlap map. Fourth, (d) we demonstrate the validation of a candidate protein within the subcompartment using immunohistochemistry and subsequent (e) co-localization analysis both within an astrocyte and within the targeted subcompartment. Fifth, we use (f) STRING analysis to assess the putative interactions and relationships among the proteins found within the subcompartment. Finally, we

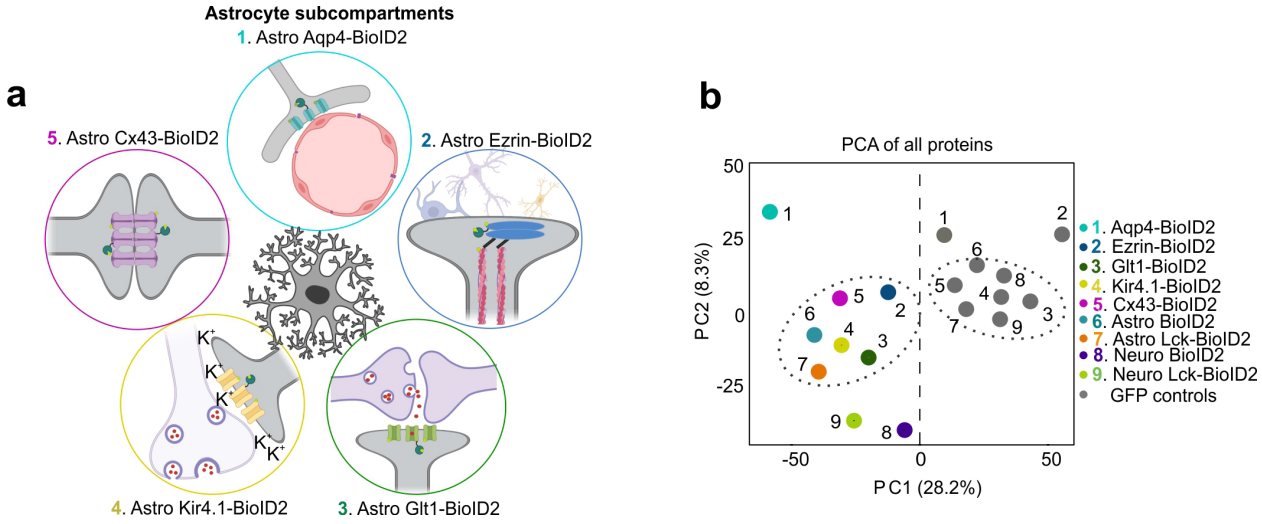

Figure 4.6: Principal component analysis of all proteins identified by mass spectrometry. a. Cartoon of the five astrocyte subcompartments targeted with BioID2. **b.** PCA plot depicting all proteins identified by mass spectrometry across 4 experimental runs per construct.

Filters	Aqp4-BioID2	Ezr-BioID2	Glt1-BioID2	Kir4.1-BioID2	Cx43-BioID2
# unique peptides	127,068				
# unique proteins	4,018	3,040	3,522	3,678	3,377
Proteins known to be contaminants and mitochondrial	3,209	2,268	2,936	3,099	2,914
# proteins with 2-fold over GFP controls and FDR <0.05	806	403	757	667	295

Table 4.1: Astrocyte BioID2 proteomics analysis filters. List of filters and resultant number of proteins used to generate high confidence lists for both astrocyte and neuron specific proteomics.

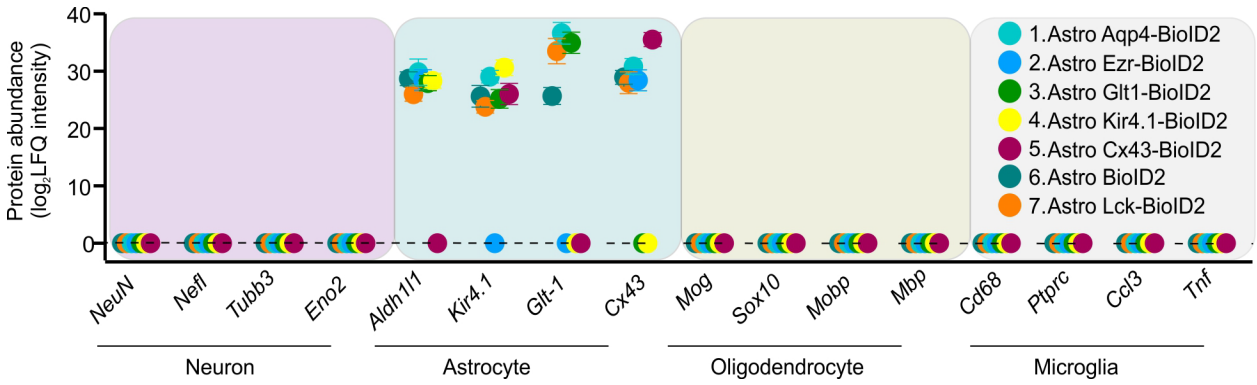

Figure 4.7: Cell-selectivity assessment of subproteome identified proteins. Expression levels (in label-free quantification intensity, LFQ intensity) of cell-specific markers identified in each BioID2 construct (n = 4 experimental runs). Proteins in this analysis represent proteins that were enriched (Log2FC > 1 versus GFP controls).

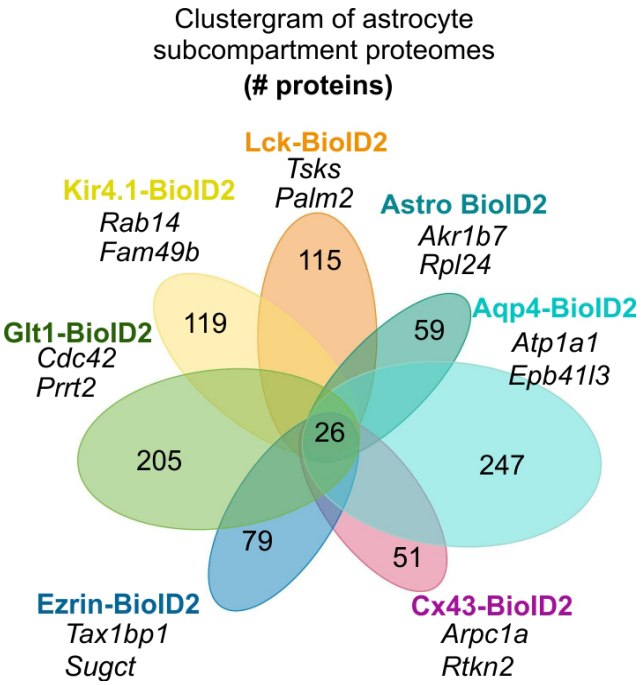

Figure 4.8: Broad assessment of cell and sub-compartment specific proteins. Clustergram shows number of unique proteins identified in each BioID2 construct experiment. Two most abundant and unique proteins for each construct are shown. The proteins represent those that were significant ($Log_2FC > 1$ and FDR < 0.05 *versus* GFP controls).

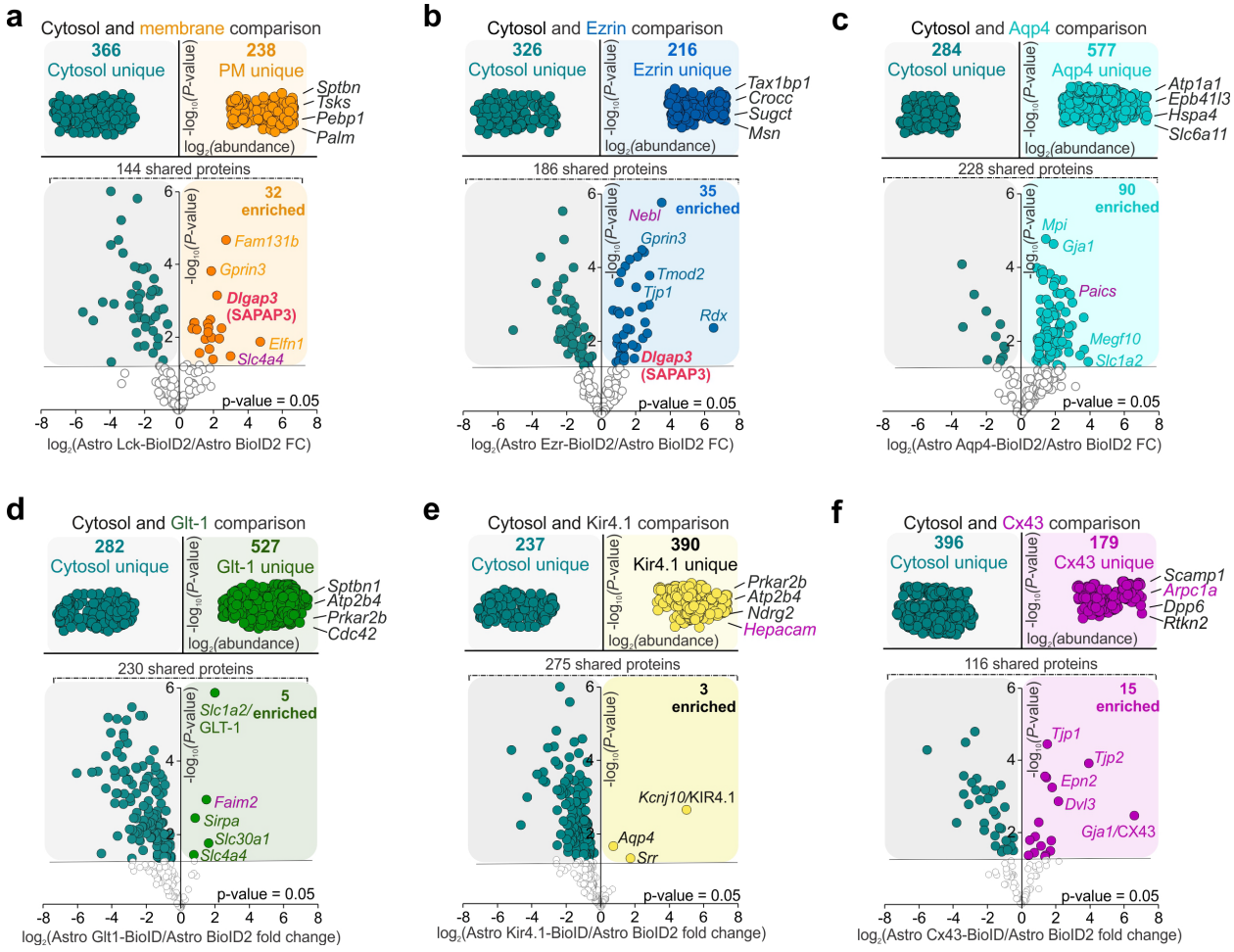

Figure 4.9: Quantitative analysis of astrocyte subproteomes. **a.** Label-free based quantification comparison of the 144 common proteins between Astro BioID2 and Astro Lck-BioID2 reveals 32 plasma membrane enriched proteins. Five highest enriched proteins ($Log_2FC > 2$) are shown. Red label shows *Dlgap3*/SAPAP3 is enriched in the membrane. Magenta label shows candidate protein chosen for expression validation **b.** As in a. but for Ezr-BioID2**. c.** as in a. but for Aqp4-BioID2. **d.** as in a, but for Glt1-BioID2 **e.** as in a. but for Kir4.1-BioID2 **f**. as in a. but for Cx43-BioID2.

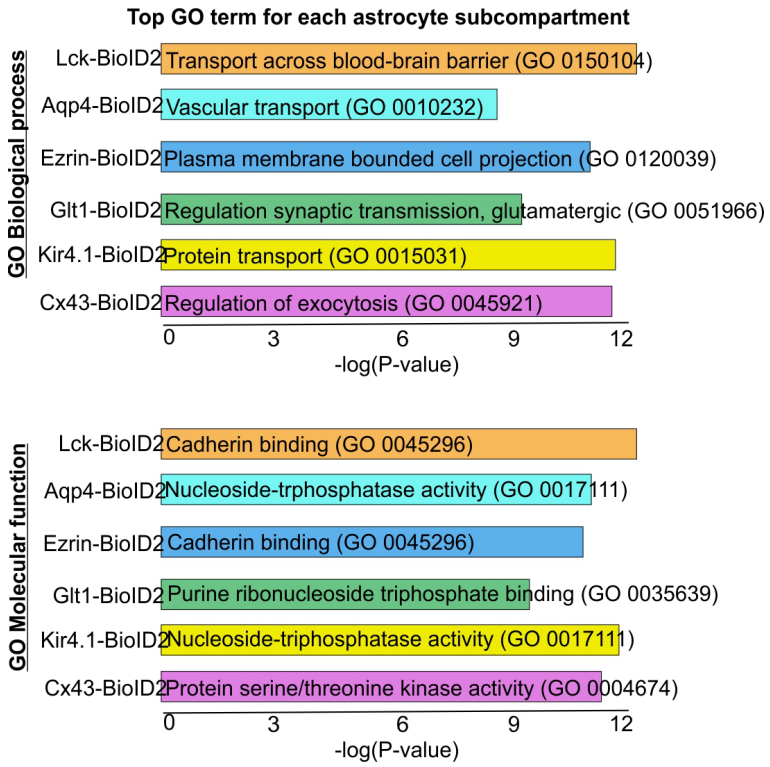

Figure 4.10: Broad functional analysis of astrocyte subproteomes. Bar graphs depict significant Enrichr gene ontology (GO) term for the unique and enriched proteins found in each astrocyte subcompartment. Top bar graph shows the GO term for "biological process," while the lower bar graph shows the GO term for "molecular function."

show the Enrichr gene ontology analysis of the proteome for "biological process," "cellular component," and "molecular function" in order to broadly assess the functions conducted by each subproteome. We created six astrocyte subproteome cards, which include all five functional subcompartments in addition to the plasma membrane. The astrocyte subproteome cards are shown as figures here, or as a compiled video as **supplemental movie 1**.

The astrocyte plasma membrane subproteome was made up of 270 proteins as shown in the quantitative comparison volcano plot when compared to cytosolic BioID2 (**Figure 4.11**). Within this list of enriched proteins, we were surprised to find SAPAP3 (gene: *Dlgap3*), a synapse associated protein. When the protein abundance (in LFQ intensity) of the 270 proteins was compared to the mRNA abundance (in FPKM) using a rank-rank hypergeometric overlap map, there was very low correlation between mRNA abundance and protein abundance. The highest correlation between both read-outs occurred with proteins that were ranked in the middle of the ranked list. Thus, the plasma membrane proteome abundance cannot be predicted with the overall astrocyte mRNA abundance. We next validated a protein that was enriched in the astrocyte plasma membrane using immunohistochemistry. We validated NBCe1 (gene: *Slc4a4*), an electrogenic sodium bicarbonate transporter. We expressed both cytosolic tdTomato red fluorescent protein and plasma membrane targeted Lck-GFP in striatal astrocytes using astrocyte-selective AAVs (*gfaABC$_1$D* tdTomato and *gfaABC$_1$D* Lck-GFP). We conducted an immunohistochemical experiment for tdTomato, Lck-GFP, and NBCe1 in the striatum where we quantified the co-localization correlation co-efficient for tdTomato and NBCe1 and Lck-GFP and NBCe1. We found that NBCe1 is expressed within astrocytes by co-localization with tdTomato (average normalized $r = 0.6$) and that it has a higher co-localization correlation co-efficient with the plasma membrane by measurement of its overlap with the Lck-GFP signal (average normalized $r = 0.8$). To further

understand the astrocyte plasma membrane proteome, we conducted a STRING analysis of putative interactions within the subproteome. We found that the proteins within the plasma membrane subproteome are highly interactive with each other (STRING protein-protein enrichment p-value = 1.89×10^{-15}) and represent specialized functions at the astrocyte plasma membrane. For example, the most significant Enrichr gene ontology term for "biological function" was "transport across the blood-brain barrier," the most significant term for "cellular component" was "adherens junctions" and thhe top molecular function term was "cadherin binding." These Enrichr terms suggest that proteins at the plasma membrane of astrocytes conduct diverse functions related to cell-cell contact and communication (adherens junctions and cadherin binding) while also regulating transport at the end foot-vasculature interface (transport across the blood brain barrier) (**Figure 4.11**). While these are plausible suggestion, careful experiments are necessary to support these claims. These experiments, however, are now possible due to this dataset.

The astrocyte Aqp4 end foot was the largest subproteome consisting of 635 proteins when compared to the cytosol. When we compared the protein abundance of these 635 proteins with their corresponding mRNA abundance using a rank-rank hypergeometric overlap map, we found that protein and mRNA abundances correlated the most in proteins of lower rank (**Figure 4.12**), thus indicating that the relationship between RNA and protein is not straightforward in the end foot. We validated the expression of end foot enriched protein, Pur6 (gene: *Paics)*, a Bifunctional phosphoribosylaminoimidazole carboxylase that is involved in de novo synthesis of second messengers including purines. We expressed both cytosolic tdTomato red fluorescent protein and end foot targeted Aqp4-GFP in striatal astrocytes using astrocyte-selective AAVs (*gfaABC$_1$D* tdTomato and *gfaABC$_1$D* Aqp4-GFP). We conducted an immunohistochemical experiment for tdTomato, Aqp4-GFP, and Pur6 in the striatum where we quantified the co-localization correlation

co-efficient for tdTomato and Pur6 and Aqp4-GFP and Pur6. We found that Pur6 is expressed within astrocytes by co-localization with tdTomato (average normalized $r = 0.5$) and that it has a higher co-localization correlation co-efficient with the astrocyte end foot by measurement of its overlap with the Aqp4-GFP signal (average normalized $r = 0.7$). We next assessed the relationship of the top 100 most abundant end foot proteins with each other by STRING analysis. STRING predicted that ~80% of the proteins at the end foot proteome interacted with one another (STRING protein-protein enrichment p-value = 1.0×10^{-16}) thus indicating a shared function of proteins at the end foot. Enrichr analysis showed that the 635 end foot proteins were significantly involved in "vascular transport" and nucleoside "triphosphosphatase activity," while the "cellular component" analysis revealed that the end foot proteome interacts with the "endoplasmic reticulum membrane." These Enrichr analysis results are exciting because they confirm that the functions of the end foot are highly relevant to its underlying structure that consists of intimate interactions with the endoplasmic reticulum (Boulay et al., 2017; Salmon et al., 2023). Furthermore, these results reveal that the astrocyte end foot proteome is complex and plays roles conducting diverse functions including energy production and subcellular transport of metabolites (**Figure 4.12**).

When we compared the astrocyte Ezrin fine processes to the cytosol, we found 234 high confidence unique and enriched proteins that make up this subproteome. Surprisingly, we found that post-synaptic associated protein, SAPAP3 was within the astrocyte fine processes. The proteome comparison with mRNA abundance via a rank-rank hypergeometric overlap map showed very little correlation between the protein abundance and the mRNA abundance. If any correlation was significant, it occurred at proteins that were relatively highly ranked (**Figure 4.13**). We validated the expression of Ezrin fine process enriched protein, Nebulette (gene: *Nebl)*, a known regulator of the actin cytoskeleton. We expressed both cytosolic tdTomato red fluorescent

protein and process targeted Ezr-GFP in striatal astrocytes using astrocyte-selective AAVs (*gfaABC$_1$D* tdTomato and *gfaABC$_1$D* Ezr-GFP). We conducted an immunohistochemical experiment for tdTomato, Ezr-GFP, and Nebulette in the striatum where we quantified the co-localization correlation co-efficient for tdTomato and Nebulette and Ezr-GFP and Nebulette. We found that Nebulette is expressed within astrocytes by co-localization with the tdTomato territory (average normalized $r = 0.6$) and that it has a higher co-localization correlation co-efficient with the astrocyte Ezrin fine processes by measurement of its overlap with the Ezr-GFP signal (average normalized $r = 0.8$) indicating that both Nebulette and Ezrin are in the astrocyte fine process subcompartments where filamentous actin is regulated. To further interrogate this assessment, we conducted a STRING analysis of putative interactions for all 234 proteins within the astrocyte Ezrin proteome. This analysis demonstrated a strong interaction within the proteome (STRING protein-protein enrichment p-value = 1.0×10^{-16}) indicating shared functions and interactions in the finest astrocyte processes. Enrichr analysis further confirmed the role Ezrin in astrocytes to regulate the actin cytoskeleton. The most significant gene ontology term for "biological process" was "plasma membrane bounded cell projection" and the most significant terms for "molecular function" were "cadherin binding" and "actin binding." Furthermore, the top two cellular component terms for the Ezrin fine process proteome was "cytoskeleton" and "actin cytoskeleton." Thus, we predict that the astrocyte fine process subproteome consists of proteins that orchestrate the regulation of filamentous actin at the most distal processes and can provide clues about astrocyte morphology. While only an assumption, we provide a list of 234 candidate proteins that may directly modulate the actin cytoskeleton and affect astrocyte association with other cells (**Figure 4.13**).

The astrocyte Glt-1 subproteome was the second largest consisting of 532 proteins when compared to the astrocyte cytosol. Comparison of the abundance of these proteins with their respective mRNA abundance using the rank-rank hypergeometric overlap map revealed that there is some correlation of mRNA and protein when the proteins were highly ranked both at the mRNA and protein levels (**Figure 4.14**). We validated a protein that was enriched in the Glt-1 subcompartment, Lifeguard2 (gene: *Faim2*), a transmembrane signaling protein that negatively regulates apoptotic mechanisms in non-dividing cells. We expressed both cytosolic tdTomato red fluorescent protein and Glt1-GFP in striatal astrocytes using astrocyte-selective AAVs (*gfaABC$_1$D* tdTomato and *gfaABC$_1$D* Glt1-GFP). We then used immunohistochemistry to assess tdTomato, Glt1-GFP, and Lifeguard2 in the striatum where we quantified the co-localization correlation co-efficient for tdTomato and Lifeguard2 and Glt1-GFP and Lifeguard2. Lifeguard2 is expressed within astrocytes by co-localization with the tdTomato territory (average normalized $r = 0.5$) and has a higher co-localization correlation co-efficient with the astrocyte Glt-1 subcompartment by measurement of its overlap with the Glt1-GFP signal (average normalized $r = 0.7$) indicating that Lifeguard2 may spatially interact with Glt-1 at the subcompartment that regulates glutamate within astrocytes. We assessed the relationship of the top 100 most abundant Glt1 subcompartment proteins with each other by STRING analysis. STRING predicted a strong association of Glt-1 associated proteins with each other (STRING protein-protein enrichment p-value = 1.0×10^{-16}) indicating functional interactions of these proteins with each other which may predict their functional output. Enrichr gene ontology analysis indeed revealed that the astrocyte Glt-1 subproteome represents proteins involved in "regulation of synaptic transmission, glutamatergic" and "modulation of chemical synaptic transmission" when we looked at functions within "biological processes." This was consistent with the role of Glt-1 at synapses which is to transport

and regulate glutamate in the extracellular space at glutamatergic synapses. When we looked at the most significant Enrichr terms for "cellular component" and "molecular function," we found "cytoplasmic vesicle membrane" and "purine ribonucleoside triphosphate binding" for each GO process, respectively. This was interesting as Glt-1 localization and expression at the plasma membrane is heavily regulated by Rab/GTPases (Martinez-Villareal et al., 2012). Thus indicating that a major function of the Glt-1 subproteome is to both regulate glutamate by transporters such as Glt-1 and also maintain and regulate Glt-1 at the plasma membrane (**Figure 4.14**).

The astrocyte Kir4.1 subproteome consisted of 393 proteins when compared to the cytosol. Among the most enriched proteins were Kir4.1 itself and its known interacting partner, Aqp4. Comparison of protein abundance to mRNA abundance of the Kir4.1 subproteome revealed that the highest correlation between these two measurements occurred where proteins were middle-ranked, and transcripts were highly ranked, again indicating the inconsistencies between transcript abundance and protein abundance (**Figure 4.15**). We next validated the protein expression Glialcam (gene: *Hepacam*) a transmembrane adhesion molecule with immunohistochemistry. We expressed both cytosolic tdTomato red fluorescent protein and Kir4.1-GFP in striatal astrocytes using astrocyte-selective AAVs (*gfaABC$_1$D* tdTomato and *gfaABC$_1$D* GFP-Kir4.1). We then used immunohistochemistry to assess tdTomato, Kir4.1-GFP, and Glialcam in the striatum where we quantified the co-localization correlation co-efficient for tdTomato and Glialcam and Kir4.1-GFP and Glialcam. Glialcam is expressed within astrocytes by co-localization with the tdTomato territory (average normalized $r = 0.7$) and has a higher co-localization correlation co-efficient with the astrocyte Kir4.1 subcompartment by measurement of its overlap with the Kir4.1-GFP signal (average normalized $r = 0.9$) indicating that Glialcam may spatially interact with Kir4.1 at the subcompartment that regulates potassium within astrocytes. To assess any putative interactions

within the Kir4.1 subproteome, we conducted STRING analysis of the 100 most abundant Kir4.1 subcompartment proteins. About 40% of these proteins were predicted to interact with one another (STRING protein-protein enrichment p-value = 1.97×10^{-6}), among these was a predicted interaction between Kir4.1, Aqp4, and Glialcam. We broadly assessed the functions of the Kir4.1 proteome with Enrichr analysis and found that the most significant gene ontology terms were "protein transport" for "biological process," "focal adhesion" for "cellular component" and again, "nucleoside triphosphatase activity" for "molecular function." For Kir4.1 function, nucleoside triphosphatase activity was unsurprising as Kir4.1 is dependent on adenosine triphosphate (ATP) for channel function (Hibino et al., 1997). However, it would be interesting to explore the function of the Kir4.1 subproteome in relation to both protein transport and focal adhesion as regulatory mechanisms of Kir4.1 and its interactors have not yet been thoroughly explored in these contexts (**Figure 4.15**).

Finally, we assessed the Cx43 astrocyte gap junction subcompartment where astrocyte interact with one another. The Cx43 gap junction subproteome consisted of 196 total unique and enriched proteins when compared to the astrocyte cytosol. Among these were Cx43 itself and tight junction scaffolding proteins ZO-1 (gene: *Tjp1*) and ZO-2 (gene: *Tjp2*). When we compared the protein abundances of the gap junction subproteome with their corresponding mRNA abundances using the rank-rank hypergeometric overlap map, we found that if any correlation occurs, it was most likely to occur to proteins that were in the middle-low rank (**Figure 4.16**). This underscores that mRNA abundances in FPKM cannot accurately predict protein abundances from mass spectrometry. We validated the expression of Arpc1a (gene: *Arpc1a*), an actin interacting protein that regulates actin polymerization and protein scaffolds. We expressed both cytosolic tdTomato red fluorescent protein and Cx43-GFP in striatal astrocytes using astrocyte-selective AAVs

(*gfaABC$_1$D* tdTomato and *gfaABC$_1$D* Cx43-GFP). We then used immunohistochemistry to assess tdTomato, Cx43-GFP, and Arpc1a in the striatum where we quantified the co-localization correlation co-efficient for tdTomato and Arpc1a and Cx43-GFP and Arpc1a. Arpc1a is expressed within astrocytes by co-localization with the tdTomato territory (average normalized $r = 0.6$) and has a higher co-localization correlation co-efficient with the astrocyte Cx43 gap junction subcompartment by measurement of its overlap with the Cx43-GFP signal (average normalized $r = 0.9$) indicating that Arpc1a may spatially interact with Cx43 at the subcompartment that is responsible for astrocyte-astrocyte contacts. To assess whether proteins found within the Cx43 gap junction subcompartment may interact with one another, we conducted STRING putative interaction analysis. This analysis revealed that there were significant interactions within the Cx43 subproteome (STRING protein-protein enrichment p-value = 3.69×10^{-12}), including a major putative connectivity hub driven by interactions with Cx43 (gene: *Gja1*). Within this connectivity hub were tight junction proteins, ZO-1 (gene: *Tjp1*) and ZO-2 (gene: *Tjp2*) which have been heavily described in the nervous system as being essential components of endothelial cell tight junctions (Corada et al., 1999). However, interactions between ZO proteins and Cx43 have been extensively validated in cell culture and cardiac studies (Giepmans and Moolenaar, 1998; Barker et al., 2002). Furthermore, in the brain, ZO-1 interactions with Cx43 gap junctional plaques have also been characterized with immunohistochemistry and co-immunoprecipitation (Penes et al., 2005). While the functional basis of Cx43 and ZO protein interactions in astrocytes remains to be explored, our data predict that regulation and localization of Cx43 astrocyte-astrocyte gap junctions may be related to the same mechanisms that orchestrate tight junction formation in other cell types. This is an exciting possibility that could be explored with more careful experiments to assess the basic biology of astrocyte gap junctions. We broadly assessed the functions of the Cx43

subproteome with Enrichr analysis and found that the most significant gene ontology terms were "regulation of exocytosis" for "biological process," "cell-cell junction" for "cellular component" and "protein serine/threonine kinase activity" for "molecular function." For Cx43 function, it was interesting to see "protein serine/threonine kinase activity" as a highly represented function as this pathway is responsible for a variety of different phosphorylation signaling events that give rise to signal transduction in a cell. The identity and function of these pathways in relation to Cx43 gap junctions remain to be explored (**Figure 4.16**).

Section 4.6: Summary and discussion

The brain functions in a coordinated effort between multiple brain regions, multiple circuits, and multiple cell types. In order to understand brain function, there is a need to understand the molecular mechanisms that give rise to multicellular interactions. As such, we have developed tailor-made tools to assess astrocyte subproteomes *in vivo.* These astrocyte subproteomes define the molecular basis for the diverse yet specialized functions conducted by these morphologically complex cells at points of interaction with other cells. We found that many subsets of proteins were preferentially localized and distributed in select astrocyte proteomes, and thus represent distinct molecular mechanisms that could be relevant for understanding astrocyte responses in different contexts such as neurodegeneration, stroke, trauma or addiction. Our data also confirmed that the relationship between gene and protein expression is complicated throughout the astrocyte subcompartments indicating the necessity to comprehensively assess the broad protein profiles of complex cells. Thus, the astrocyte subproteomes we present here provide a bounty of information including novel proteins and signaling pathways that remain to be deeply explored in astrocytes. One of these proteins was post-synapse associated protein, SAPAP3 (gene: *Dlgap3*).

Canonically, SAPAP3 has been described as a post-synaptic density associated protein that functions as a scaffold to anchor glutamate receptors to the post-synaptic membrane in conjunction with other proteins at the glutamatergic synapse (Rasmussen et al., 2017). Surprisingly, using astrocyte-selective and subcompartment specific BioID2, we found SAPAP3 expressed reliably at the astrocyte plasma membrane and at the Ezrin fine processes, indicating SAPAP3 localization at subcompartments where astrocytes contact neurons. Our discovery that SAPAP3 is expressed in striatal astrocytes is novel and underscores the exciting discoveries that have yet to be described upon further exploration of the astrocyte subproteomes presented here.

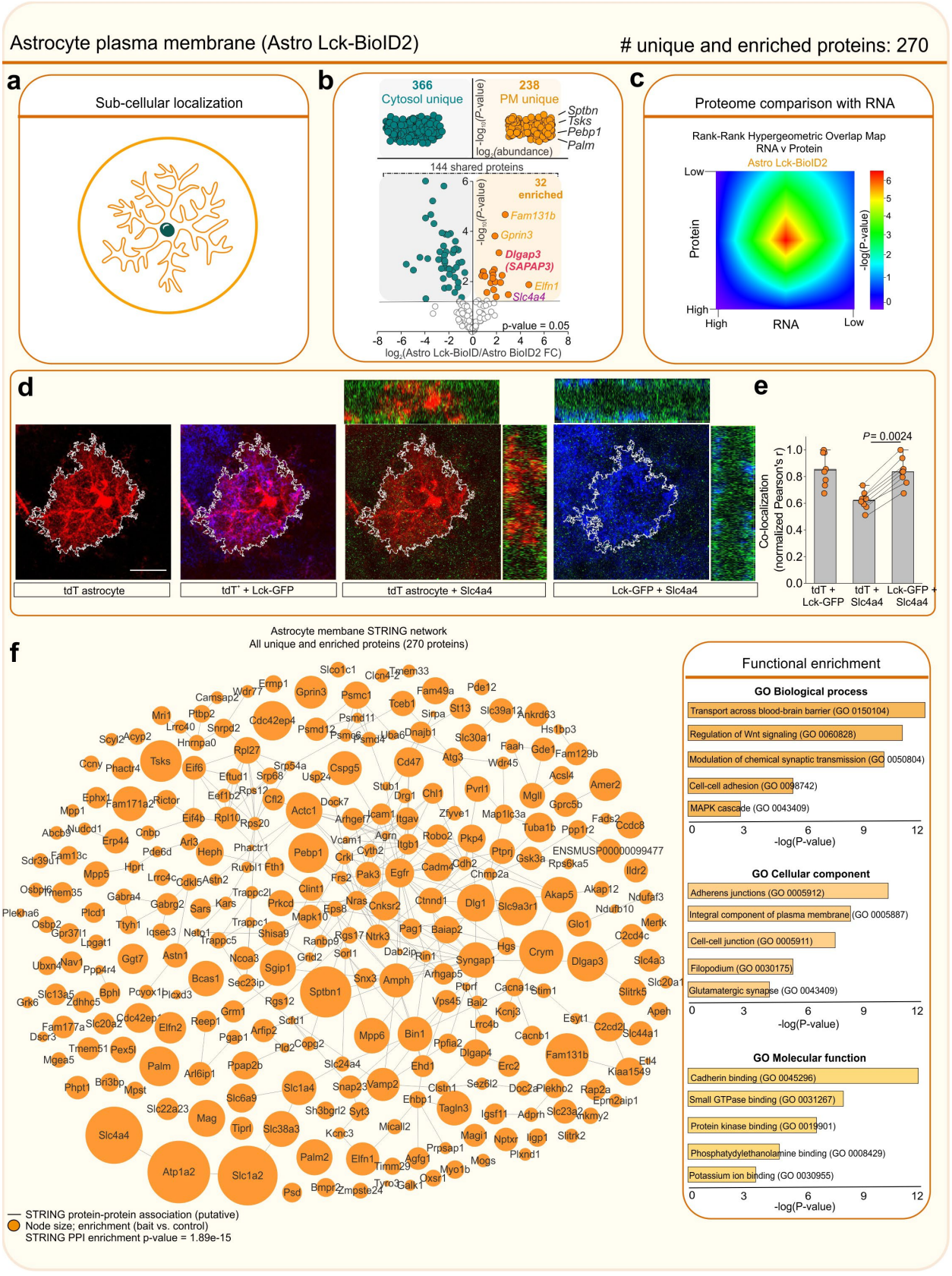
Astrocyte plasma membrane (Astro Lck-BioID2)
unique and enriched proteins: 270
a
Sub-cellular localization
b
366
Cytosol unique
238
PM unique
Sptbn
Tsks
Pebp1
Palm
-log10(P-value)
log2(abundance)
144 shared proteins
32
enriched
Fam131b
Gprin3
Dlgap3
(SAPAP3)
Elfn1
Slc4a4
p-value = 0.05
log2(Astro Lck-BioID/Astro BioID2 FC)
c
Proteome comparison with RNA
Rank-Rank Hypergeometric Overlap Map
RNA v Protein
Astro Lck-BioID2
Low
High
Protein
RNA
-log(P-value)
d
tdT astrocyte
tdT+ Lck-GFP
tdT astrocyte + Slc4a4
Lck-GFP + Slc4a4
e
P= 0.0024
Co-localization
(normalized Pearson's r)
tdT +
Lck-GFP
tdT +
Slc4a4
Lck-GFP +
Slc4a4
f
Astrocyte membane STRING network
All unique and enriched proteins (270 proteins)
STRING protein-protein association (putative)
Node size; enrichment (bait vs. control)
STRING PPI enrichment p-value = 1.89e-15
Functional enrichment
GO Biological process
Transport across blood-brain barrier (GO 0150104)
Regulation of Wnt signaling (GO 0060828)
Modulation of chemical synaptic transmission (GO 0050804)
Cell-cell adhesion (GO 0098742)
MAPK cascade (GO 0043409)
-log(P-value)
GO Cellular component
Adherens junctions (GO 0005912)
Integral component of plasma membrane (GO 0005887)
Cell-cell junction (GO 0005911)
Filopodium (GO 0030175)
Glutamatergic synapse (GO 0043409)
-log(P-value)
GO Molecular function
Cadherin binding (GO 0045296)
Small GTPase binding (GO 0031267)
Protein kinase binding (GO 0019901)
Phosphatydylethanolamine binding (GO 0008429)
Potassium ion binding (GO 0030955)
-log(P-value)

Figure 4.11: Astrocyte Card 1: plasma membrane. a. BioID2 that is targeted to the plasma membrane biotinylates proteins at the plasma membrane. **b**. Venn diagram depicts enriched and significant protein identifications ($Log_2FC > 1$ and FDR < 0.05 versus GFP controls) overlapped between astrocyte cytosol (Astro BioID2) and astrocyte plasma membrane (Astro Lck-BioID2). The top five most abundant unique proteins for the plasma membrane are shown. Label-free based quantification comparison of the 144 common proteins between Astro BioID2 and Astro Lck-BioID2 reveals 32 plasma membrane enriched proteins. Five highest enriched proteins ($Log_2FC > 2$) are shown. Red label shows *Dlgap3*/SAPAP3 is enriched in the membrane. Magenta label shows candidate protein chosen for expression validation. **c**. Heat map shows the rank-rank hypergeometric overlap (RRHO) of the RNA and protein rank for the 270 plasma membrane proteins. Each pixel represents the significance of overlap between the two datasets in $-\log_{10}$(P-value). Red pixels represent highly significant overlap. Color scale denotes the range of P-values at the negative $\log_{10}$ scale (Bin size = 100). **d**. Immunohistochemical analysis of Slc4a4 (Nbc1) protein in tdTomato and Lck-GFP labeled astrocytes shows co-localization within the astrocyte territory. **e**. Co-localization analysis using Pearson's r co-efficient shows high co-localization between Lck-GFP and Slc4a4 (Nbc1) (n = 8 tdT+ cells; Paired t-test). **f**. Scale-free STRING analysis interaction map of the 270 unique and enriched biotinylated proteins identified in astrocyte plasma membrane with Astro Lck-BioID2 . Node size represents the enrichment of each protein vs the GFP control (log2(BioID2/GFP)). Edges represent known interactions from the STRING database. Bar graphs show the functional enrichment analysis of the 270 proteins using "Biological process", "Cellular component", and "Molecular function" terms from Enrichr.

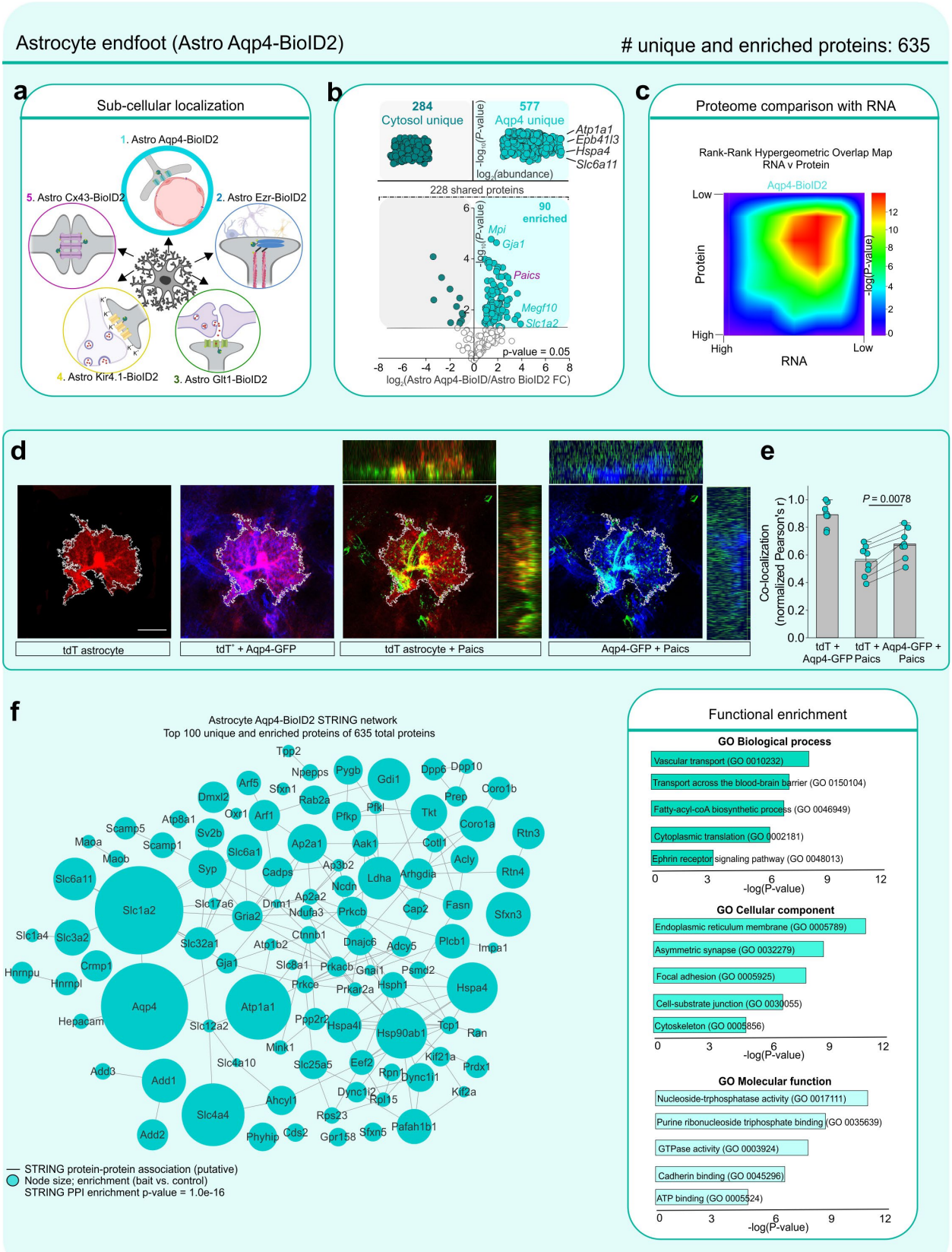
Astrocyte endfoot (Astro Aqp4-BioID2)
unique and enriched proteins: 635
a
Sub-cellular localization
1. Astro Aqp4-BioID2
2. Astro Ezr-BioID2
3. Astro Glt1-BioID2
4. Astro Kir4.1-BioID2
5. Astro Cx43-BioID2
b
284
Cytosol unique
577
Aqp4 unique
Atp1a1
Epb41l3
Hspa4
Slc6a11
228 shared proteins
90
enriched
Mpi
Gja1
Paics
Megf10
Slc1a2
p-value = 0.05
c
Proteome comparison with RNA
Rank-Rank Hypergeometric Overlap Map
RNA v Protein
Aqp4-BioID2
Protein
RNA
Low
High
d
tdT astrocyte
tdT+ + Aqp4-GFP
tdT astrocyte + Paics
Aqp4-GFP + Paics
e
P = 0.0078
Co-localization
(normalized Pearson's r)
tdT + Aqp4-GFP
tdT + Paics
Aqp4-GFP + Paics
f
Astrocyte Aqp4-BioID2 STRING network
Top 100 unique and enriched proteins of 635 total proteins
STRING protein-protein association (putative)
Node size; enrichment (bait vs. control)
STRING PPI enrichment p-value = 1.0e-16
Functional enrichment
GO Biological process
Vascular transport (GO 0010232)
Transport across the blood-brain barrier (GO 0150104)
Fatty-acyl-coA biosynthetic process (GO 0046949)
Cytoplasmic translation (GO 0002181)
Ephrin receptor signaling pathway (GO 0048013)
GO Cellular component
Endoplasmic reticulum membrane (GO 0005789)
Asymmetric synapse (GO 0032279)
Focal adhesion (GO 0005925)
Cell-substrate junction (GO 0030055)
Cytoskeleton (GO 0005856)
GO Molecular function
Nucleoside-triphosphatase activity (GO 0017111)
Purine ribonucleoside triphosphate binding (GO 0035639)
GTPase activity (GO 0003924)
Cadherin binding (GO 0045296)
ATP binding (GO 0005524)
-log(P-value)

Figure 4.12: Astrocyte Card 2: endfoot. a. BioID2 that is targeted with Aqp4-BioID2 biotinylates proteins at the astrocyte endfoot. **b**. Venn diagram depicts enriched and significant protein identifications ($Log_2FC > 1$ and FDR < 0.05 versus GFP controls) overlapped between astrocyte cytosol (Astro BioID2) and the astrocyte endfoot (Astro Aqp4-BioID2). The top five most abundant unique proteins for the endfoot are shown. Label-free based quantification comparison of the 228 common proteins between Astro BioID2 and Astro Aqp4-BioID2 reveals 58 endfoot enriched proteins. Five highest enriched proteins ($Log_2FC > 2$) are shown. Magenta label shows candidate protein chosen for expression validation. **c**. Heat map shows the rank-rank hypergeometric overlap (RRHO) of the RNA and protein rank for the 635 endfoot proteins. Each pixel represents the significance of overlap between the two datasets in $-\log_{10}$(P-value). Red pixels represent highly significant overlap. Color scale denotes the range of P-values at the negative $\log_{10}$ scale (Bin size = 100). **d**. Immunohistochemical analysis of PAICS protein in tdTomato and Lck-GFP labeled astrocytes shows co-localization within the astrocyte territory.

e. Co-localization analysis using Pearson's r co-efficient shows high co-localization between Aqp4-GFP and PAICS (n = 8 tdT+ cells; Paired t-test). **f**. Scale-free STRING analysis interaction map of the top 100 unique and enriched biotinylated proteins identified in the astrocyte endfoot with Astro Aqp4-BioID2 . Node size represents the enrichment of each protein vs the GFP control (log2(BioID2/GFP)). Edges represent known interactions from the STRING database. Bar graphs show the functional enrichment analysis of the 635 proteins using "Biological process", "Cellular component", and "Molecular function" terms from Enrichr.

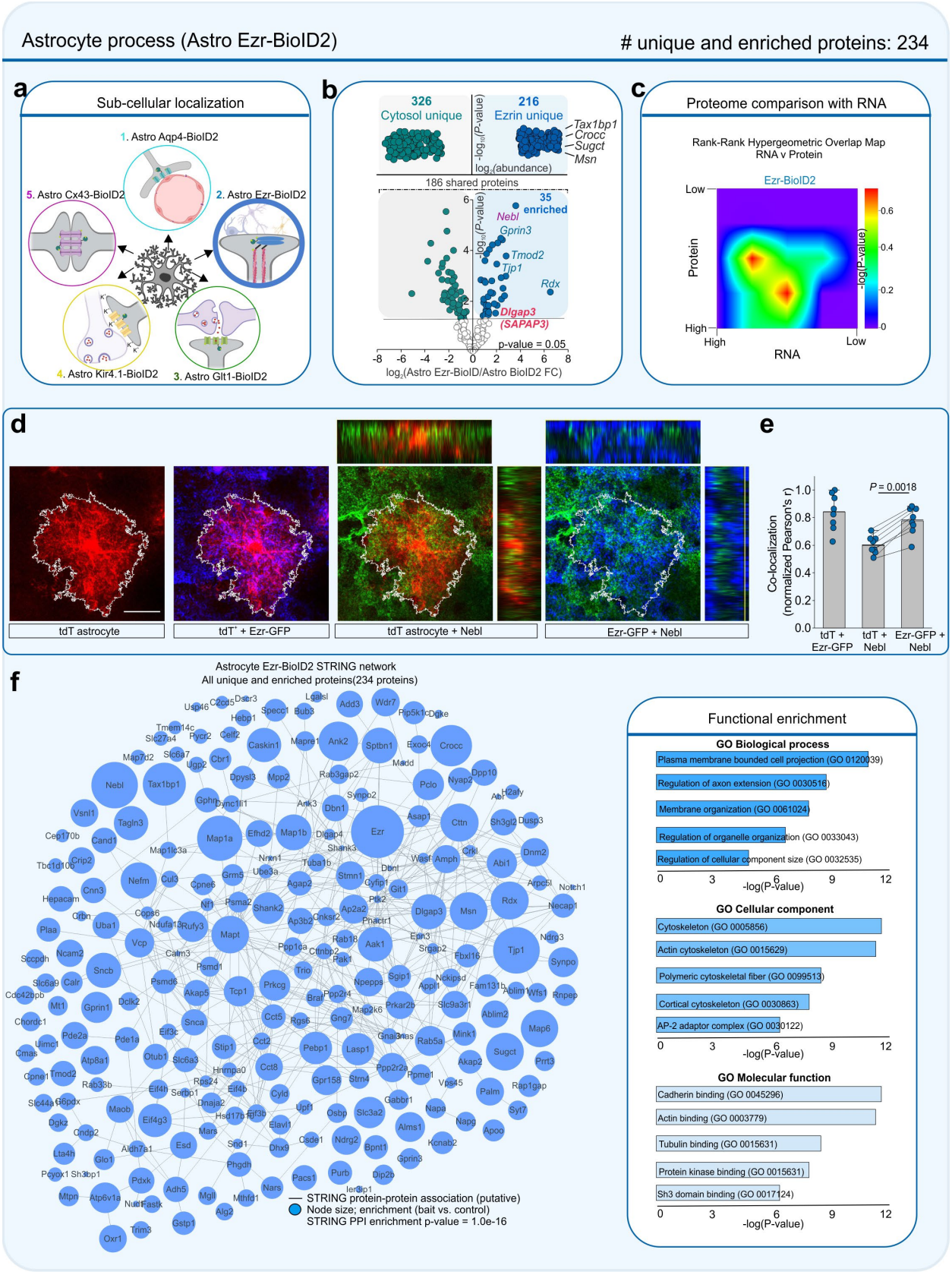
Astrocyte process (Astro Ezr-BioID2)
unique and enriched proteins: 234
a
Sub-cellular localization
1. Astro Aqp4-BioID2
2. Astro Ezr-BioID2
3. Astro Glt1-BioID2
4. Astro Kir4.1-BioID2
5. Astro Cx43-BioID2
b
326
Cytosol unique
216
Ezrin unique
Tax1bp1
Crocc
Sugct
Msn
186 shared proteins
35
enriched
Nebl
Gprin3
Tmod2
Tjp1
Rdx
Dlgap3
(SAPAP3)
p-value = 0.05
log2(Astro Ezr-BioID/Astro BioID2 FC)
c
Proteome comparison with RNA
Rank-Rank Hypergeometric Overlap Map
RNA v Protein
Ezr-BioID2
Protein
RNA
Low
High
d
tdT astrocyte
tdT+ Ezr-GFP
tdT astrocyte + Nebl
Ezr-GFP + Nebl
e
P = 0.0018
Co-localization
(normalized Pearson's r)
tdT + Ezr-GFP
tdT + Nebl
Ezr-GFP + Nebl
f
Astrocyte Ezr-BioID2 STRING network
All unique and enriched proteins(234 proteins)
STRING protein-protein association (putative)
Node size; enrichment (bait vs. control)
STRING PPI enrichment p-value = 1.0e-16
Functional enrichment
GO Biological process
Plasma membrane bounded cell projection (GO 0120039)
Regulation of axon extension (GO 0030516)
Membrane organization (GO 0061024)
Regulation of organelle organization (GO 0033043)
Regulation of cellular component size (GO 0032535)
-log(P-value)
GO Cellular component
Cytoskeleton (GO 0005856)
Actin cytoskeleton (GO 0015629)
Polymeric cytoskeletal fiber (GO 0099513)
Cortical cytoskeleton (GO 0030863)
AP-2 adaptor complex (GO 0030122)
GO Molecular function
Cadherin binding (GO 0045296)
Actin binding (GO 0003779)
Tubulin binding (GO 0015631)
Protein kinase binding (GO 0015631)
Sh3 domain binding (GO 0017124)

Figure 4.13: Astrocyte Card 3: astrocyte process. a. BioID2 that is targeted with Ezr-BioID2 biotinylates proteins at astrocyte processes. **b**. Venn diagram depicts enriched and significant protein identifications ($Log_2FC > 1$ and FDR < 0.05 versus GFP controls) overlapped between astrocyte cytosol (Astro BioID2) and Astro Ezr-BioID2. The top five most abundant unique proteins for Astro Ezr-BioID2 are shown. Label-free based quantification comparison of the 186 common proteins between Astro BioID2 and Astro Ezr-BioID2 reveals 32 process enriched proteins. Five highest enriched proteins ($Log_2FC > 2$) are shown. Magenta label shows candidate protein chosen for expression validation. **c**. Heat map shows the rank-rank hypergeometric overlap (RRHO) of the RNA and protein rank for the 234 Ezr-BioID2 proteins. Each pixel represents the significance of overlap between the two datasets in $-\log_{10}$(P-value). Red pixels represent highly significant overlap. Color scale denotes the range of P-values at the negative $\log_{10}$ scale (Bin size = 100). **d**. Immunohistochemical analysis of Nebl protein in tdTomato and Ezr-GFP labeled astrocytes shows co-localization within the astrocyte territory.**e**. Co-localization analysis using Pearson's r co-efficient shows high co-localization between Ezr-GFP and Nebl (n = 8 tdT+ cells; Paired t-test). **f**. Scale-free STRING analysis interaction map of the 234 unique and enriched biotinylated proteins identified in astrocyte processes with Astro Ezr-BioID2 . Node size represents the enrichment of each protein vs the GFP control (log2(BioID2/GFP)). Edges represent known interactions from the STRING database. Bar graphs show the functional enrichment analysis of the 234 proteins using "Biological process", "Cellular component", and "Molecular function" terms from Enrichr.

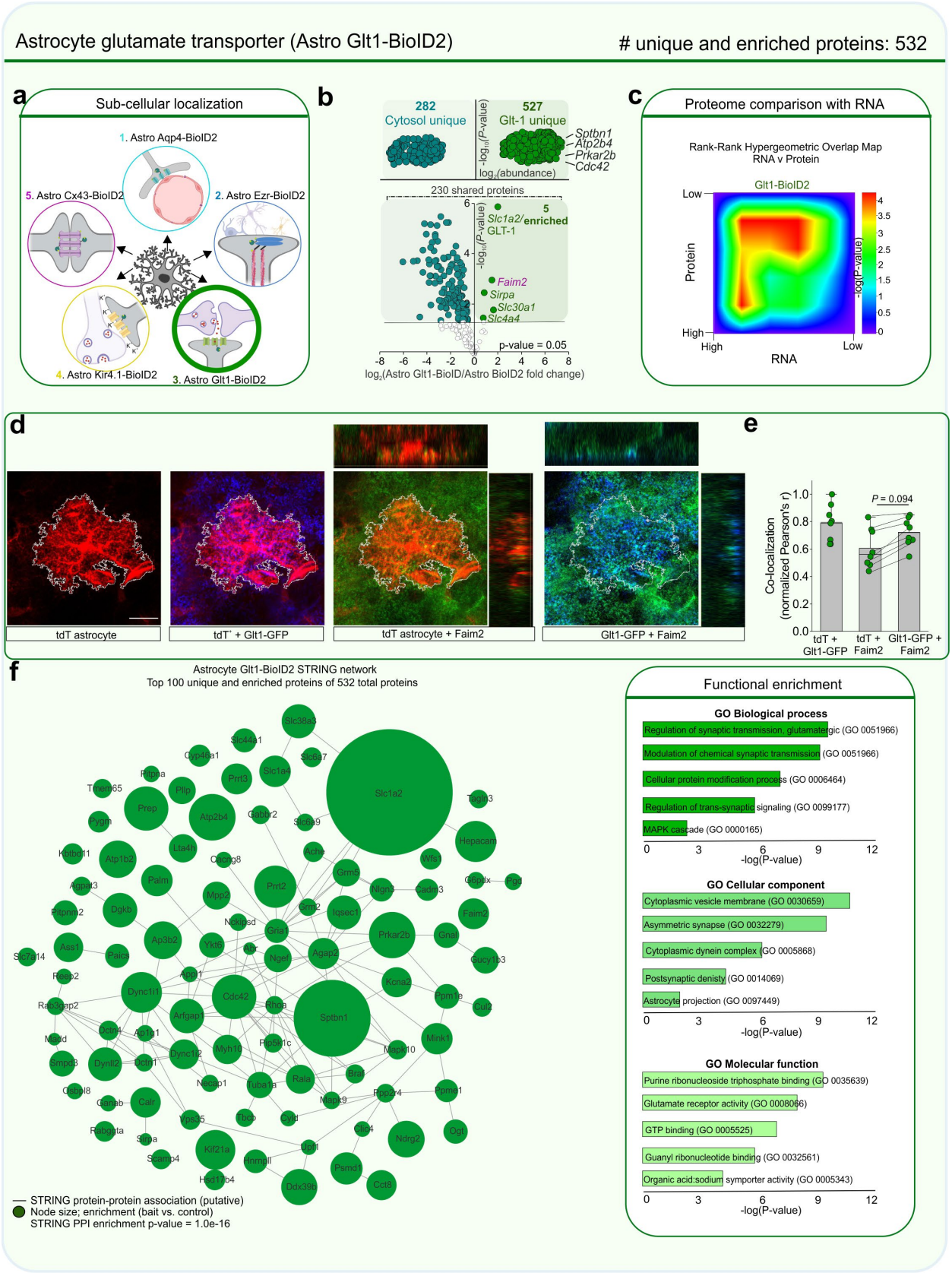
Astrocyte glutamate transporter (Astro Glt1-BioID2)
unique and enriched proteins: 532
a
Sub-cellular localization
1. Astro Aqp4-BioID2
2. Astro Ezr-BioID2
3. Astro Glt1-BioID2
4. Astro Kir4.1-BioID2
5. Astro Cx43-BioID2
b
282
Cytosol unique
527
Glt-1 unique
Sptbn1
Atp2b4
Prkar2b
Cdc42
-log10(P-value)
log2(abundance)
230 shared proteins
5
Slc1a2/enriched
GLT-1
Faim2
Sirpa
Slc30a1
Slc4a4
p-value = 0.05
log2(Astro Glt1-BioID/Astro BioID2 fold change)
c
Proteome comparison with RNA
Rank-Rank Hypergeometric Overlap Map
RNA v Protein
Glt1-BioID2
Protein
RNA
Low
High
-log(P-value)
d
tdT astrocyte
tdT+ + Glt1-GFP
tdT astrocyte + Faim2
Glt1-GFP + Faim2
e
Co-localization
(normalized Pearson's r)
P = 0.094
tdT + Glt1-GFP
tdT + Faim2
Glt1-GFP + Faim2
f
Astrocyte Glt1-BioID2 STRING network
Top 100 unique and enriched proteins of 532 total proteins
STRING protein-protein association (putative)
Node size; enrichment (bait vs. control)
STRING PPI enrichment p-value = 1.0e-16
Functional enrichment
GO Biological process
Regulation of synaptic transmission, glutamatergic (GO 0051966)
Modulation of chemical synaptic transmission (GO 0051966)
Cellular protein modification process (GO 0006464)
Regulation of trans-synaptic signaling (GO 0099177)
MAPK cascade (GO 0000165)
GO Cellular component
Cytoplasmic vesicle membrane (GO 0030659)
Asymmetric synapse (GO 0032279)
Cytoplasmic dynein complex (GO 0005868)
Postsynaptic denisty (GO 0014069)
Astrocyte projection (GO 0097449)
GO Molecular function
Purine ribonucleoside triphosphate binding (GO 0035639)
Glutamate receptor activity (GO 0008066)
GTP binding (GO 0005525)
Guanyl ribonucleotide binding (GO 0032561)
Organic acid:sodium symporter activity (GO 0005343)
-log(P-value)

Figure 4.14: Astrocyte Card 4: astrocyte glutamate transporter. a. BioID2 that is targeted with Glt1-BioID2 biotinylates proteins at astrocyte sites of glutamate uptake. **b**. Venn diagram depicts enriched and significant protein identifications ($Log_2FC > 1$ and $FDR < 0.05$ versus GFP controls) overlapped between astrocyte cytosol (Astro BioID2) and Astro Glt1-BioID2. The top five most abundant unique proteins for Astro Glt1-BioID2 are shown. Label-free based quantification comparison of the 230 common proteins between Astro BioID2 and Astro Glt1-BioID2 reveals 5 enriched proteins. Magenta label shows candidate protein chosen for expression validation. **c**. Heat map shows the rank-rank hypergeometric overlap (RRHO) of the RNA and protein rank for the 532 Glt1-BioID2 proteins. Each pixel represents the significance of overlap between the two datasets in $-\log_{10}$(P-value). Red pixels represent highly significant overlap. Color scale denotes the range of P-values at the negative $\log_{10}$ scale (Bin size = 100). **d**. Immunohistochemical analysis of Faim2 protein in tdTomato and Glt1-GFP labeled astrocytes shows co-localization within the astrocyte territory. **e**. Co-localization analysis using Pearson's r co-efficient shows high co-localization between Glt1-GFP and Faim2 (n = 8 tdT+ cells; Paired t-test). **f**. Scale-free STRING analysis interaction map of the top 100 unique and enriched biotinylated proteins identified with Astro Glt1-BioID2 . Node size represents the enrichment of each protein vs the GFP control (log2(BioID2/GFP)). Edges represent known interactions from the STRING database. Bar graphs show the functional enrichment analysis of all 532 proteins using "Biological process", "Cellular component", and "Molecular function" terms from Enrichr.

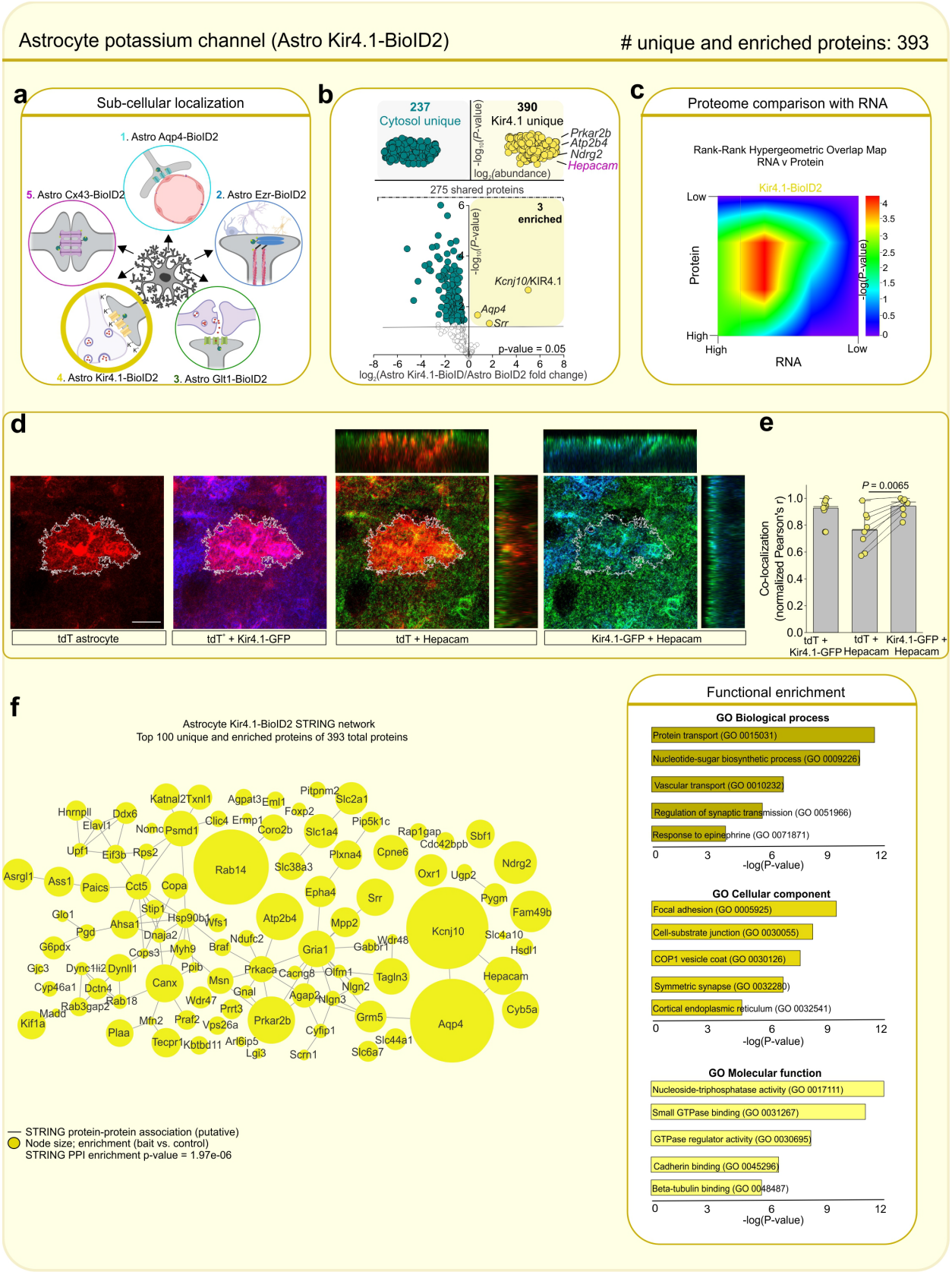
Astrocyte potassium channel (Astro Kir4.1-BioID2)
unique and enriched proteins: 393
a
Sub-cellular localization
1. Astro Aqp4-BioID2
2. Astro Ezr-BioID2
3. Astro Glt1-BioID2
4. Astro Kir4.1-BioID2
5. Astro Cx43-BioID2
b
237
Cytosol unique
390
Kir4.1 unique
Prkar2b
Atp2b4
Ndrg2
Hepacam
275 shared proteins
3
enriched
Kcnj10/KIR4.1
Aqp4
Srr
p-value = 0.05
log2(Astro Kir4.1-BioID/Astro BioID2 fold change)
c
Proteome comparison with RNA
Rank-Rank Hypergeometric Overlap Map
RNA v Protein
Kir4.1-BioID2
Protein
RNA
d
tdT astrocyte
tdT+ + Kir4.1-GFP
tdT + Hepacam
Kir4.1-GFP + Hepacam
e
P = 0.0065
Co-localization (normalized Pearson's r)
tdT + Kir4.1-GFP
tdT + Hepacam
Kir4.1-GFP + Hepacam
f
Astrocyte Kir4.1-BioID2 STRING network
Top 100 unique and enriched proteins of 393 total proteins
STRING protein-protein association (putative)
Node size; enrichment (bait vs. control)
STRING PPI enrichment p-value = 1.97e-06
Functional enrichment
GO Biological process
Protein transport (GO 0015031)
Nucleotide-sugar biosynthetic process (GO 0009226)
Vascular transport (GO 0010232)
Regulation of synaptic transmission (GO 0051966)
Response to epinephrine (GO 0071871)
-log(P-value)
GO Cellular component
Focal adhesion (GO 0005925)
Cell-substrate junction (GO 0030055)
COP1 vesicle coat (GO 0030126)
Symmetric synapse (GO 0032280)
Cortical endoplasmic reticulum (GO 0032541)
GO Molecular function
Nucleoside-triphosphatase activity (GO 0017111)
Small GTPase binding (GO 0031267)
GTPase regulator activity (GO 0030695)
Cadherin binding (GO 0045296)
Beta-tubulin binding (GO 0048487)

Figure 4.15: Astrocyte Card 5: astrocyte potassium channel. a. BioID2 that is targeted with Kir4.1-BioID2 biotinylates proteins at astrocyte sites of potassium uptake. **b**. Venn diagram depicts enriched and significant protein identifications ($Log_2FC > 1$ and $FDR < 0.05$ versus GFP controls) overlapped between astrocyte cytosol (Astro BioID2) and Astro Kir4.1-BioID2. The top five most abundant unique proteins for Astro Kir4.1-BioID2 are shown. Label-free based quantification comparison of the 275 common proteins between Astro BioID2 and Astro Kir4.1-BioID2 reveals 3 enriched proteins. Magenta label shows candidate protein chosen for expression validation. **c**. Heat map shows the rank-rank hypergeometric overlap (RRHO) of the RNA and protein rank for the 393 Kir4.1-BioID2 proteins. Each pixel represents the significance of overlap between the two datasets in $-\log_{10}$(P-value). Red pixels represent highly significant overlap. Color scale denotes the range of P-values at the negative $\log_{10}$ scale (Bin size = 100). **d**. Immunohistochemical analysis of Hepacam protein in tdTomato and Kir4.1-GFP labeled astrocytes shows co-localization within the astrocyte territory. **e**. Co-localization analysis using Pearson's r co-efficient shows high co-localization between Kir4.1-GFP and Hepacam (n = 8 tdT+ cells; Paired t-test). **f**. Scale-free STRING analysis interaction map of the top 100 unique and enriched biotinylated proteins identified with Astro Kir4.1-BioID2 . Node size represents the enrichment of each protein vs the GFP control (log2(BioID2/GFP)). Edges represent known interactions from the STRING database. Bar graphs show the functional enrichment analysis of all 393 proteins using "Biological process", "Cellular component", and "Molecular function" terms from Enrichr.

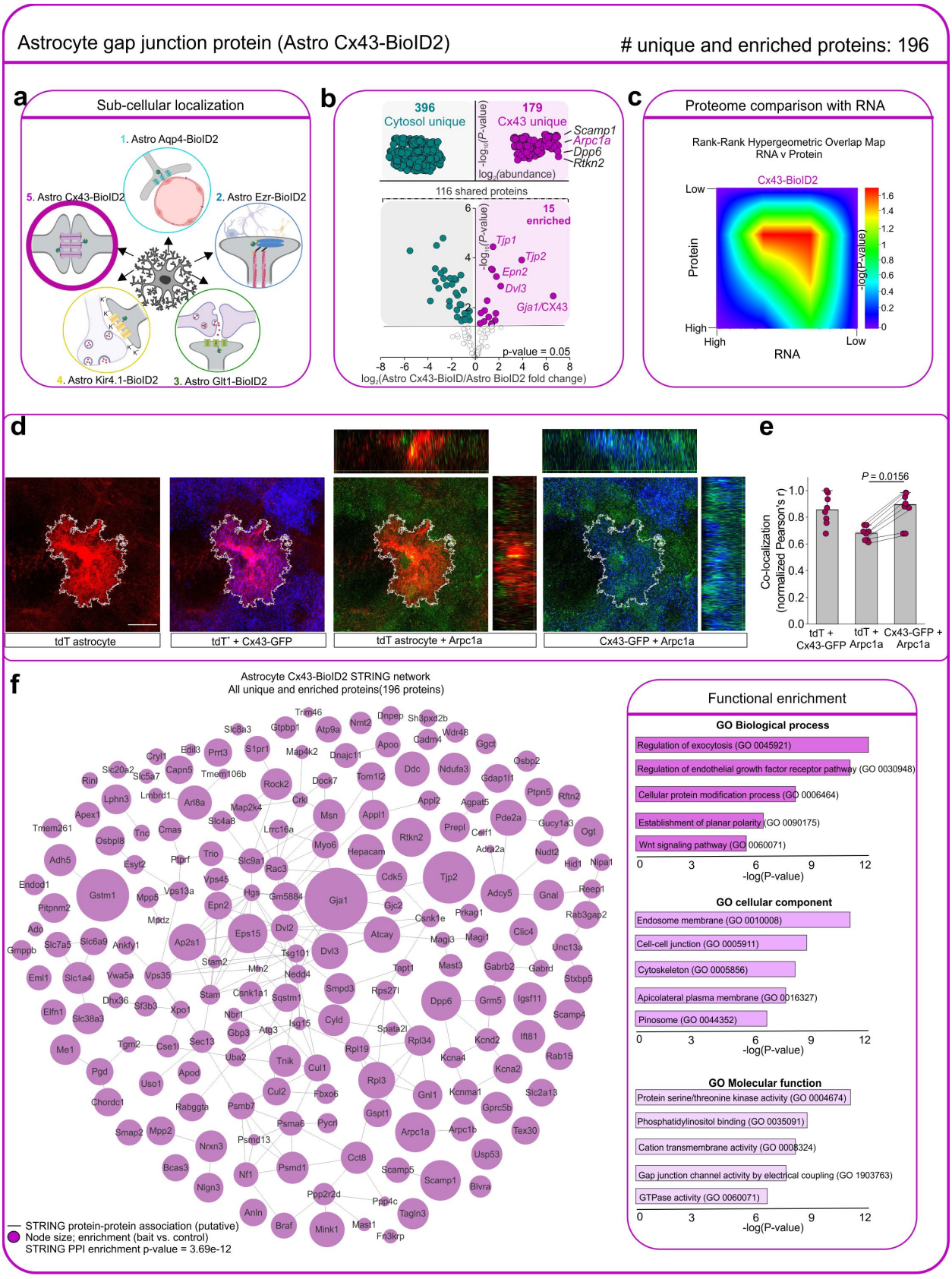
Astrocyte gap junction protein (Astro Cx43-BioID2)
unique and enriched proteins: 196
a
Sub-cellular localization
1. Astro Aqp4-BioID2
2. Astro Ezr-BioID2
3. Astro Glt1-BioID2
4. Astro Kir4.1-BioID2
5. Astro Cx43-BioID2
b
396
Cytosol unique
179
Cx43 unique
Scamp1
Arpc1a
Dpp6
Rtkn2
-log10(P-value)
log2(abundance)
116 shared proteins
15
enriched
Tjp1
Tjp2
Epn2
Dvl3
Gja1/CX43
p-value = 0.05
log2(Astro Cx43-BioID/Astro BioID2 fold change)
c
Proteome comparison with RNA
Rank-Rank Hypergeometric Overlap Map
RNA v Protein
Cx43-BioID2
Low
High
Protein
RNA
-log(P-value)
d
tdT astrocyte
tdT+ + Cx43-GFP
tdT astrocyte + Arpc1a
Cx43-GFP + Arpc1a
e
P = 0.0156
Co-localization
(normalized Pearson's r)
tdT + Cx43-GFP
tdT + Arpc1a
Cx43-GFP + Arpc1a
f
Astrocyte Cx43-BioID2 STRING network
All unique and enriched proteins(196 proteins)
STRING protein-protein association (putative)
Node size; enrichment (bait vs. control)
STRING PPI enrichment p-value = 3.69e-12
Functional enrichment
GO Biological process
Regulation of exocytosis (GO 0045921)
Regulation of endothelial growth factor receptor pathway (GO 0030948)
Cellular protein modification process (GO 0006464)
Establishment of planar polarity (GO 0090175)
Wnt signaling pathway (GO 0060071)
-log(P-value)
GO cellular component
Endosome membrane (GO 0010008)
Cell-cell junction (GO 0005911)
Cytoskeleton (GO 0005856)
Apicolateral plasma membrane (GO 0016327)
Pinosome (GO 0044352)
GO Molecular function
Protein serine/threonine kinase activity (GO 0004674)
Phosphatidylinositol binding (GO 0035091)
Cation transmembrane activity (GO 0008324)
Gap junction channel activity by electrical coupling (GO 1903763)
GTPase activity (GO 0060071)

Figure 4.16: Astrocyte Card 6: astrocyte gap junction protein. a. BioID2 that is targeted with Cx43-BioID2 biotinylates proteins at astrocyte sites of gap junction coupling. **b**. Venn diagram depicts enriched and significant protein identifications ($Log_2FC > 1$ and $FDR < 0.05$ versus GFP controls) overlapped between astrocyte cytosol (Astro BioID2) and Astro Cx43-BioID2. The top five most abundant unique proteins for Astro Cx43-BioID2 are shown. Label-free based quantification comparison of the 116 common proteins between Astro BioID2 and Astro Cx43-BioID2 reveals 17 enriched proteins. Magenta label shows candidate protein chosen for expression validation. **c**. Heat map shows the rank-rank hypergeometric overlap (RRHO) of the RNA and protein rank for the 196 Cx43-BioID2 proteins. Each pixel represents the significance of overlap between the two datasets in $-\log_{10}$(P-value). Red pixels represent highly significant overlap. Color scale denotes the range of P-values at the negative $\log_{10}$ scale (Bin size = 100). **d**. Immunohistochemical analysis of Arpc1a protein in tdTomato and Cx43-GFP labeled astrocytes shows co-localization within the astrocyte territory. **e**. Co-localization analysis using Pearson's r co-efficient shows high co-localization between Cx43-GFP and Aprc1a (n = 8 tdT+ cells; Paired t-test). **f**. Scale-free STRING analysis interaction map of the 196 unique and enriched biotinylated proteins identified with Astro Cx43-BioID2 . Node size represents the enrichment of each protein vs the GFP control (log2(BioID2/GFP)). Edges represent known interactions from the STRING database. Bar graphs show the functional enrichment analysis of all 196 proteins using "Biological process", "Cellular component", and "Molecular function" terms from Enrichr.

CHAPTER 5: MOLECULAR MECHANISMS OF SAPAP3 IN ASTROCYTES

Section 5.1: Introduction

While transmembrane receptors and ion channels are the primary method by which cells respond to information from other cells, these specialized receptors do not function in isolation from other proteins within a cell. Transmembrane signaling receptors and ion channels are heavily regulated, localized, and maintained in discrete subcellular domains by protein complexes composed of a variety of different proteins that maintain the intracellular signaling milieu. A relevant example of these mechanisms occurs at the post-synaptic density (PSD) where a post-synaptic neuron integrates information from a pre-synaptic neuron during neurotransmission.

In the mammalian central nervous system about 60% of synapses are excitatory and use glutamate as a neurotransmitter (Nadler, 2012). Excitatory synaptic neurotransmission occurs primarily at contacts between presynaptic boutons, and actin-rich protrusions on the post-synaptic neuron called dendritic spines which are enveloped by a post-synaptic plasma membrane. During neurotransmission, the molecular composition of post-synaptic membranes and the chemical modifications of synaptic proteins are essential determinants of the strength and connectivity of excitatory synapses (Scannevin and Huganir 2000; Sheng and Kim 2011). The post-synaptic density (PSD) is a thick, electron-dense structure that is directly adjacent to the post-synaptic membrane and is composed of a variety of proteins including cytoskeletal tethered receptor ion channels, signaling molecules, adhesion proteins, cytoskeletal proteins, and most notably, protein complexes composed of scaffolding proteins (Sheng and Hoogenraad 2007). While, arguably, all proteins composing the PSD are crucial for synaptic function, scaffolding proteins are the interface by which all components of the PSD assemble into a specialized macromolecular complex that is specially localized, stabilized, and regulated at the post-synaptic membrane. A conserved and key

protein family involved in this scaffolding activity is the SAPAP (genes: *Dlgap 1-4*) family of proteins (Rasmussen et al., 2017).

While four major SAPAP isoforms are found to be expressed in the mammalian brain, SAP90/PSD-95-associated protein isoform 3, also known as SAPAP3 (gene: *Dlgap3*) is expressed in the striatum, cortex, thalamus, hippocampus, and cerebellum (Welch et al., 2004). Comparison within these five regions shows that SAPAP3 is preferentially enriched and highest expressed in the striatum (Welch et al., 2004). The full length structure of SAPAP3, is characterized by a repeat domain composed of a 14-amino-acid repeats, a dynein light-chain domain, three proline-rich domains, and a GKAP homology (GH1) domain, all by which SAPAP3 binds to its interacting partners in neurons (Kim et al., 1997; Takeuchi et al., 1997; Naisbitt et al., 2000). At the PSD, these binding domains allow SAPAP3 to couple both ionotropic and metabotropic glutamate receptors to the actin cytoskeleton via actin binding molecules, such as Cortactin (**Figure 5.1**), and thus link multiple cellular signaling pathways together (Chen et al., 2000; Kim and Shang 2004; Bats et al., 2007). Furthermore, SAPAP3 can also directly interact with signaling molecules such as Nos and ArgBPs to modulate both cell adhesion and the PSD-adjacent actin cytoskeleton (Cestra et al., 2005; Lee et al., 2016). Because of its ability to bind a plethora of different proteins due to its diverse binding domains, SAPAP3 presents itself as an important scaffolding protein that may regulate other signaling pathways or interact with other proteins independent of the PSD.

Given that SAPAP3 plays a well-characterized and crucial function at glutamatergic synapses, it was surprising to discover that SAPAP3 could reliably be identified in the cytosol, plasma membrane, and fine processes of striatal astrocytes with cell and subcompartment specific proteomic methods (**Chapter 3**, **Chapter 4, Figure 5.2**). While no study has conclusively stated

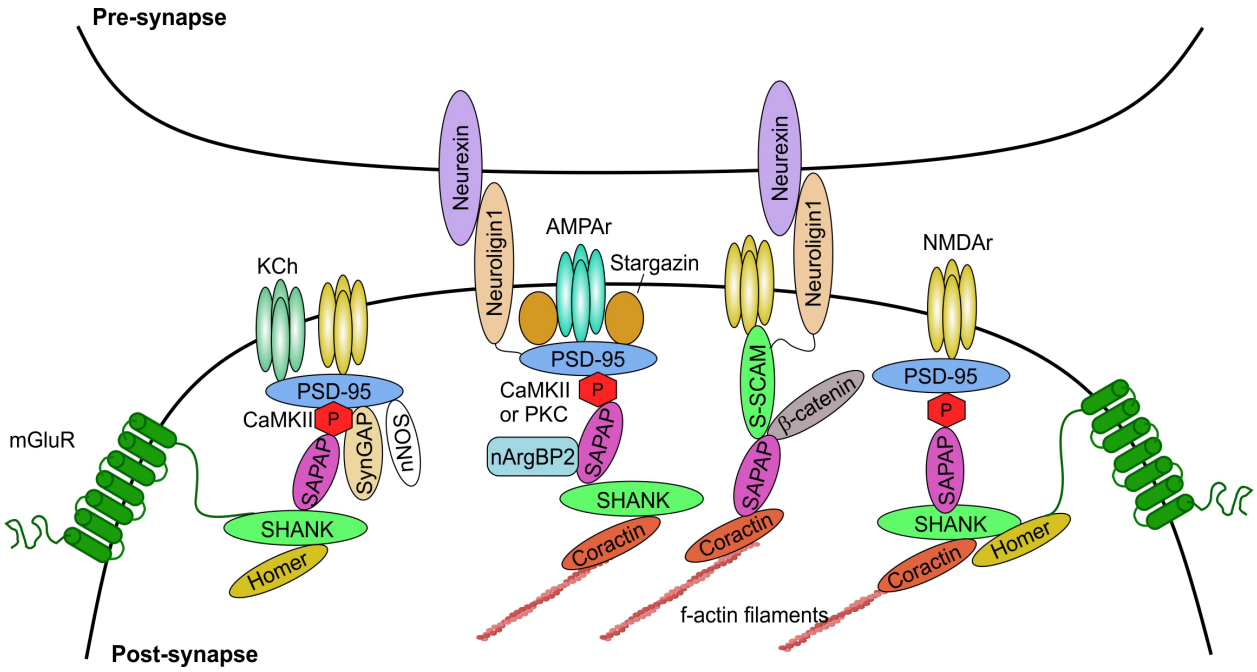

Figure 5.1: SAPAP3 interactions at the postsynaptic density. Cartoon depicting canonical SAPAP3 interactors at neuronal postsynaptic densities of glutamatergic synapses.

that SAPAP3 is unique to neurons, we sought to confirm and validate that SAPAP3 is indeed expressed in striatal astrocytes. Using a variety of approaches including deep RNA-sequencing, single-cell RNA-sequencing, *in situ* hybridization, and immunohistochemistry, we confirmed that SAPAP3 is expressed within striatal astrocytes. Furthermore, we used molecular biology approaches to dissect the mechanisms by which SAPAP3 functions within astrocytes. Our goal was to assess how SAPAP3 signaling in astrocytes compares to the known SAPAP3 signaling mechanisms in neurons. In this way, we have taken a protein that was previously undescribed in astrocytes and assessed how it regulates astrocyte function in a local circuit where this protein is known to be enriched.

Section 5.2: Multi-omic evidence of SAPAP3 expression in striatal astrocytes

Using our cell and subcompartment specific proteomics methods, we discovered that SAPAP3 (gene: *Dlgap3*) is reliably detected in the cytosol, plasma membrane, and fine processes (Ezr-BioID2) of striatal astrocytes (**Figure 5.2**). Additionally, quantitative analysis revealed that SAPAP3 is equivalently expressed in both the astrocyte and neuronal plasma membranes at the protein level **(Chapter 3** and **Figure 5.2**). Because a protein product does require a transcript to be expressed, we assessed the expression of SAPAP3 at the RNA level within astrocytes and neurons using cell-specific *in vivo* RiboTag deep RNA-sequencing. Assessment of the SAPAP3 mRNA abundance (in FPKM) revealed that although SAPAP3 mRNA abundance is indeed higher in neurons (FPKM of 110) than astrocytes (FPKM of 60), astrocytes still express high levels of SAPAP3 mRNA transcripts (**Figure 5.3**). For context, a gene is considered highly expressed if the FPKM value is larger than 10 (Trapnell et al., 2010). Therefore, at the mRNA level, astrocytes of the striatum indeed express SAPAP3. When we compared the RiboTag immunoprecipitated (IP)

mRNA from astrocytes and neurons to the mRNA from the soup of all cells (input fraction), we found that SAPAP3 was not enriched in either astrocytes or neurons ($log_2FC < 1$ IP vs input). However, canonical astrocyte enriched genes such as *Slc1a2* and *Aldh1l1* were enriched in astrocytes compared to input ($Log_2FC > 1$ for both). Likewise, canonical neuron enriched genes such as *Rbfox3* and *Ppp1r1b* were enriched compared to input ($Log_2FC > 1$ for both). This result indicated that in addition to expression within neurons and astrocytes, SAPAP3 may be expressed in other brain cell types in the striatum (**Figure 5.3**).

In order to compare the transcript levels of SAPAP3 to the protein levels detected by mass spectrometry, we ranked all the mRNA transcripts found in neurons and astrocytes from highest abundance (rank of 1) to lowest abundance; we also ranked all the neuron and astrocyte specific proteins from highest abundance (rank of 1) to lowest abundance. When we assessed the rank of SAPAP3 at the mRNA level, SAPAP3 was ranked 481 in striatal neurons and ranked 960 in striatal astrocytes. At the protein level, SAPAP3 was ranked 10 in striatal neurons and ranked 20 in striatal astrocytes (**Figure 5.3**). Therefore, one of the reasons that SAPAP3 expression in astrocytes had been overlooked is because its expression is not in the top 10% of genes that are highly expressed or enriched in astrocytes. This underscores the idea that meaningful biology may be missed by filtering genes or proteins based on cell specificity and enrichment. Perhaps a plethora of interesting genes and proteins are understudied because they are expressed in multiple cell types.

We next assessed the expression of SAPAP3 in neurons and astrocytes using single-cell RNA sequencing to confirm SAPAP3 expression across three independent approaches. We utilized a striatum specific single-cell sequencing dataset (Yu et al., 2020) and found that SAPAP3 was expressed within *Gja1* positive astrocytes and *Rbfox3* positive neurons (**Figure 5.4**). Additionally, when we compared the expression level of SAPAP3 transcripts in *Gja1* positive astrocytes to

Rbfox3 positive neurons, we found that although there are less astrocytes that express SAPAP3, both neurons and astrocytes that express SAPAP3 at equivalent levels. This supports the deep sequencing results that both neurons and astrocytes express SAPAP3 at comparable levels and that striatal astrocytes do indeed express SAPAP3.

Finally, we coupled the multi-omic approaches here with RNA-scope fluorescent *in situ* hybridization (FISH) and immunohistochemistry (IHC) to confirm expression of SAPAP3 mRNA and protein within astrocytes. To do this we used both wild-type mice and SAPAP3 constitutive knockout mice (SAPAP3 KO) as a control for our probes and antibodies. Additionally, the mice were microinjected with astrocyte specific tdTomato AAV (AAV 2/5 *gfaABC$_1$D*-tdTomato) in the striatum for assessment of astrocytes and their territories. The FISH data showed that fluorescent signal of SAPAP3 mRNA was found within the tdTomato+ soma and territory indicating bona-fide expression of SAPAP3 within astrocytes in the striatum. This fluorescent signal was undetectable in tdTomato+ astrocytes of SAPAP3 KO mice (**Figure 5.5**). Furthermore, using IHC with an anti-SAPAP3 antibody, we found that SAPAP3 protein was indeed expressed within wild-type tdTomato+ astrocytes and their territories, while the protein signal was undetectable in SAPAP3 KO tdTomato+ astrocytes (**Figure 5.5**). Therefore, both our SAPAP3 mRNA and protein probes were specific to SAPAP3 and thus, SAPAP3 is expressed in striatal astrocytes.

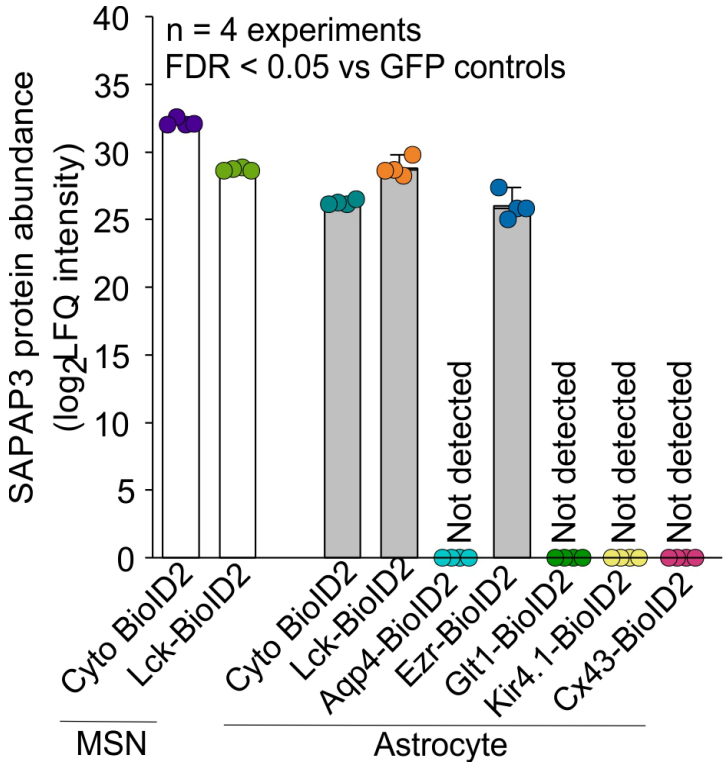

Figure 5.2: SAPAP3 protein expression across subcompartments. Bar graph depicts SAPAP3 protein abundance across the neuronal and astrocytic subcompartments from BioID2 experiments. If SAPAP3 was not detected it was labeled as "not detected" and given an abundance value of 0.

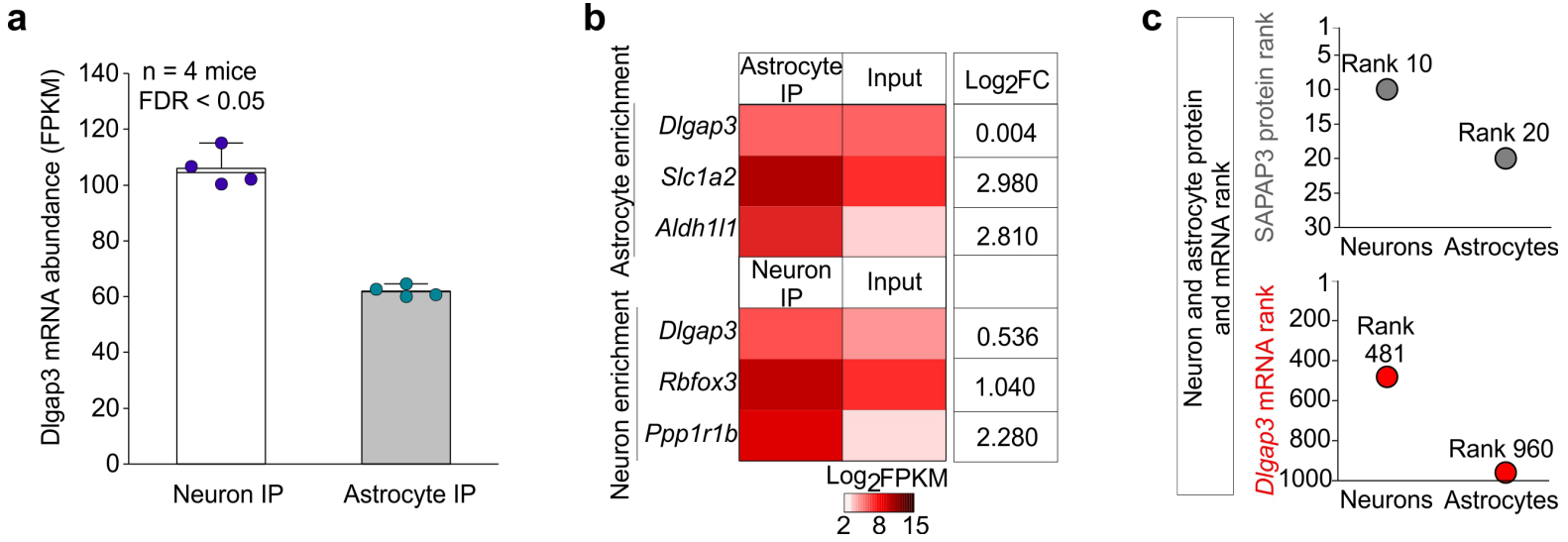

Figure 5.3: SAPAP3 mRNA expression in neurons and astrocytes. a. Bar graph depicts SAPAP3 mRNA abundance in FPKM in striatal neurons and astrocytes expressing RiboTag. IP, immunoprecipitated mRNA from Ribotag experiments. b. Heat map shows the log_2FPKM expression of SAPAP3 mRNA and select cell-enriched markers for astrocytes and neurons from the IP samples and from the input sample. c. Top graph shows the SAPAP3 protein rank from the cell-specific BioID2 proteomics experiments. Bottom graph shows the SAPAP3 mRNA rank from cell-specific RiboTag transcriptomics experiments. The rank was based on abundance for both graphs.

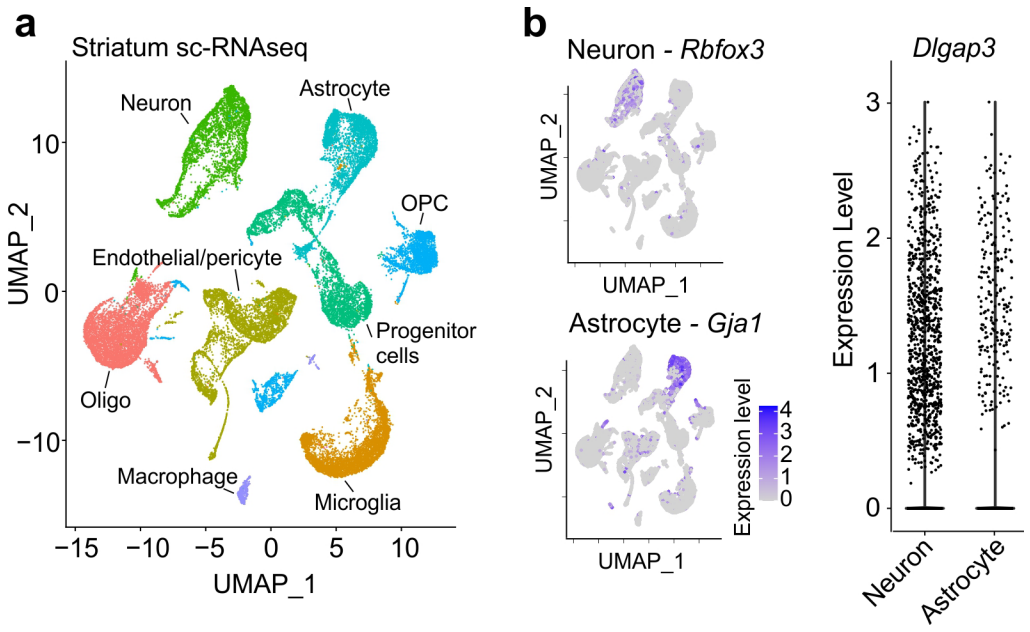

Figure 5.4: SAPAP3 mRNA expression in neurons and astrocytes at the single-cell level. **a.** Uniform manifold approximation and projection (UMAP) plot of striatal cells (n = 31,956 individual cells; replotted from published scRNA-seq (Yu *et al*., 2020)). **b.** Subpanels show expression of *Rbfox3* in neurons and *Gja1* in the astrocytes. Violin plot showing relative expression level of *Dlgap3* per cell in the neurons and astrocytes. Expression level is defined as the Log_2 normalized gene count per cell.

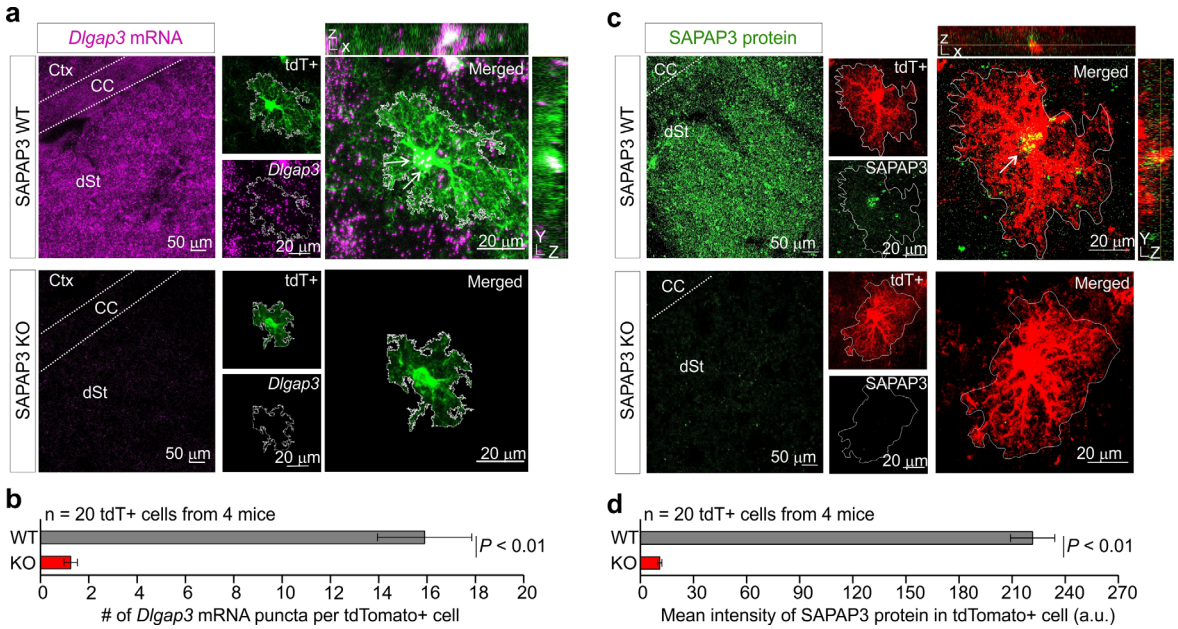

Figure 5.5: SAPAP3 mRNA and protein is expressed in striatal astrocytes. **a.** Representative image of dorsal striatum from either WT or SAPAP3 KO mice labeled by RNAscope *in situ* hybridization for *Dlgap3* mRNA (purple) and by IHC for tdTomato+ astrocytes (green). Left images show *Dlgap3* mRNA abundance throughout the entire striatum. Expanded images show dorsal striatum tdTomato+ astrocytes express *Dlgap3* mRNA (white arrows). **b.** Bar graph shows the number of *Dlgap3* mRNA puncta within tdTomato+ astrocytes in either WT or SAPAP3 KO mice. Mean and SEM are shown; n = 20 tdTomato+ astrocytes from 4 mice per group (Two-tailed Mann-Whitney test). **c.** Representative image of dorsal striatum from either WT or SAPAP3 KO mice labeled by IHC for SAPAP3 protein (green). Left images show SAPAP3 protein abundance in the striatum. Expanded images show tdTomato+ dorsal striatum astrocytes express SAPAP3 protein (white arrow). **g.** Bar graph depicting intensity in arbitrary units (a.u.) of SAPAP3 protein within tdTomato+ astrocytes in either WT or SAPAP3 KO mice. Mean and SEM are shown; n = 20 tdTomato+ astrocytes from 4 mice per group (Two-tailed Mann-Whitney test).

Section 5.3: SAPAP3 interactions in astrocytes

Using three multi-omic approaches in conjunction with *in situ* hybridization and immunohistochemistry, we show that SAPAP3 is expressed in striatal astrocytes at both the RNA and protein level. However, if this is the case, we asked whether astrocytic SAPAP3 had similar molecular interactors as those known for neuronal SAPAP3. As discussed previously, neuronal SAPAP3 is known to interact with proteins localized to the PSD where it mediates scaffolding. These conclusions arose from studies using PSD-95 (a known post-synaptic density marker protein) affinity purification mass spectrometry and co-immunoprecipitation coupled with western blot experiments. Therefore, the direct interactomes of SAPAP3 itself have not been assessed in either astrocytes or neurons.

To generate a list of high-confidence putative interactors of SAPAP3 in astrocytes and neurons, we fused BioID2 to SAPAP3 and generated cell-selective AAVs: one for neurons (AAV 1 *hSyn1* HA-BioID2-SAPAP3) and one for astrocytes (AAV 2/5 *gfaABC$_1$D* HA-BioID2-SAPAP3). As controls we developed similar AAVs where the BioID2 was replaced with a GFP (AAV 1 *hSyn1* eGFP-SAPAP3 and AAV 2/5 *gfaABC$_1$D* eGFP-SAPAP3). To validate the expression and cell-selectivity of these constructs, we microinjected each AAV construct unilaterally into the striatum. After three weeks of expression, exogenous biotin was subcutaneously injected into the mice for seven days. Sixteen hours after the last injection, the mouse brain tissues were processed for immunohistochemistry and western blot.

Immunohistochemistry analysis showed that expression of astrocyte selective BioID2-SAPAP3 and neuron selective BioID2-SAPAP3 was confined to S100β+ astrocytes and DARPP-32+ neurons, respectively (**Figure 5.6**). The analysis also demonstrated that about 75% of S100β+ astrocytes had BioID2 expression with the astrocyte selective AAV, while about 78% of DARPP+

neurons had BioID2 expression via the neuron selective AAV. Additionally, using a streptavidin-conjugated probe, biotinylated protein signal within astrocytes and neurons was visible and was proximal to the HA-BioID2 signal for each construct. For example, both the signal for astrocyte selective BioID2-SAPAP3 and its corresponding streptavidin signal looked to be bushy and localized to the plasma membrane. Similarly, the signal for neuron selective BioID2-SAPAP3 and its corresponding streptavidin signal were visualized at the soma and in the dendrite-dense neuropil. Taken together, the IHC analysis confirmed that our cell-selective BioID2-SAPAP3 constructs were expressed in their intended targets and were functional.

When we assessed the biotinylated proteins with western blotting and the streptavidin-HRP probe, both of our cell-selective BioID2-SAPAP3 constructs displayed a significant increase in biotinylation signal across the molecular weight spectra when compared to the contralateral non-injected side (**Figure 5.7**). Thus, BioID2-SAPAP3 is expressed and is functional in striatal astrocytes and neurons, *in vivo.*

Given that our astrocyte and neuron selective BioID2-SAPAP3 constructs were both cell-selective and functional, we next sought to identify the proximal biotinylated interactors of SAPAP3 in both cell types by using mass spectrometry. We injected a set of mice with either astrocyte or neuron selective BioID2-SAPAP3 AAV bilaterally into the striatum. We also injected another set of mice with the cognate cell-selecive GFP AAV control (AAV 1 *hSyn1* eGFP-SAPAP3 or AAV 2/5 *gfaABC$_1$D* eGFP-SAPAP3). After three weeks of expression, all mice received the same subcutaneous injection of 24 mg/kg of biotin for 7 days, after which the mouse striata were processed for purification of biotinylated proteins.

Using high performance liquid chromatography coupled with tandem mass spectrometry as described previously, a total of 26,783 unique peptides were run across the cell-selective samples,

including the controls. Of these, 11,363 peptides were from astrocyte selective BioID2-SAPAP3 and 15,420 were from neuron selective BioID2-SAPAP3. After filtering mitochondrial proteins and known contaminants out, there were 968 proteins in the Astro BioID2-SAPAP3 and Astro GFP-SAPAP3 combined list while there were 1,459 proteins for Neuro BioID2-SAPAP3 and Neuro GFP-SAPAP3 combined list. To generate a high-confidence list of putative interactors of SAPAP3 in each cell type, we chose proteins that were present in 3 or more mass spectrometry experiments and were statistically significant versus the GFP controls ($Log_2FC > 1$ and FDR < 0.05). With this analysis we found 170 high confidence SAPAP3 interactors in astrocytes and 423 high confidence SAPAP3 interactors in neurons (**Table 5.1**).

Using the LFQ values as in previous analyses, we then quantitatively compared the 170 high-confidence astrocyte SAPAP3 interactors with the 423 high-confidence neuron SAPAP3 interactors. When we compared both interaction lists, we found 49 unique astrocyte SAPAP3 interactors and 306 unique neuron SAPAP3 interactors. Additionally, 228 proteins were shared between both cell types with 223 enriched in neurons ($log_2FC > 1$ and FDR < 0.05 vs astrocyte) and 5 present in both cell types, including SAPAP3 itself. There were no enriched proteins in the shared protein list in astrocytes (**Figure 5.8**). Therefore, the 49 unique proteins found in astrocytes constitute the high-confidence astrocytic SAPAP3 interactome, while the 306 unique proteins in addition to the 223 enriched proteins constitute the high confidence neuronal SAPAP3 interactome.

To further understand the biology and functions related to each cell-specifc SAPAP3 interactome, we conducted a significance analysis of interactome (SAINT) to assign confidence scores to the protein-protein interactions between SAPAP3 and its putative interactors (Choi et al., 2011). In this way, we use the whole of the mass spectrometry data including the spectral counts

for both MS1 and MS2 in addition to the LFQ intensity to assess the *in silico* probability that each protein found in the high-confidence interactor list is indeed a bona-fide interactor. Additionally, we assessed the function of the SAPAP3 interactors of each cell-type using PANTHER gene ontology analysis. SAINT analysis revealed that for the 49 proteins comprising the astrocytic SAPAP3 interactome, all had a SAINT statistical score of less than 0.05 (BFDR < 0.05) and a SAINT probability score of greater than 0.5 (SAINT > 0.5) indicating that, according to the mass spectrometry data, these are the highest confidence interactors of SAPAP3 in astrocytes. When we assessed the GO molecular functions of these astrocyte specific interacting proteins we found proteins related to five main functional terms: actin cytoskeleton binding, G-protein signaling, protein localization, and glutamate signaling including astrocytic glutamate transporters Glt-1 (gene: *Slc1a2*) and GLAST (Gene: *Slc1a3*). We conducted the same analysis for the top 50 neuronal SAPAP3 interacting proteins and found that all proteins in this interactome had a SAINT statistical score of less than 0.05 (BFDR < 0.05) and a SAINT probability score of greater than 0.6 (SAINT > 0.6) indicating these are the highest confidence interactors of SAPAP3 in neurons. When we assessed the GO molecular functions of these neuron specific interacting proteins we found proteins related to two main functional terms: G-protein signaling and glutamate signaling, which included both ionotropic and metabotropic glutamate receptors, as well as vital components of the PSD such as PSD-95 (gene: *Dlg4*) and Homer (gene: *Homer*), all which are known neuronal interactors of SAPAP3 at excitatory synapses (**Figure 5.9**).

Our astrocyte and neuron-specific SAPAP3 interactome datasets revealed cell specific interactors of SAPAP3 and highlight shared and separable mechanisms of SAPAP3 in both cell types (**Figure 5.9**). In astrocytes, SAPAP3 is an interactor of actin related proteins such as Ezrin (**Chapter 4**) and glutamate homeostatic proteins, including Glt-1, GLAST, and monoamine

oxidase B which catalyzes the conversion of glutamate to GABA (gene: *Maob*). Furthermore, it interacts with proteins related to protein localization indicating that perhaps, in astrocytes, SAPAP3 also functions as a scaffolding protein to localize signaling proteins such as channels and transporters to peri-synaptic astrocyte processes and stabilizes them there via interactions with the underlying actin cytoskeleton. In neurons, we found proteins related to the PSD which have already been validated as neuronal SAPAP3 interactors (Chen et al., 2000; Kim and Shang 2004; Bats et al., 2007) as well as some related to G-protein signaling. Thus, our proteomic data define SAPAP3 interactions that are cell-specific (Glt-1 in astrocytes) and provides clues to molecular mechanisms in astrocytes and neurons that are shared (related to glutamatergic signaling) and different (protein localization).

Section 5.4: Validation of astrocytic SAPAP3 interactors

In astrocytes, the major interactions of SAPAP3 were related to glutamate homeostasis and the actin cytoskeleton. To assess these key interactions, we first validated direct protein-protein interactions between SAPAP3 and Glt-1 (gene: *Slc1a2*) and between SAPAP3 and Ezrin (gene: *Ezr*). Co-immunoprecipitation (co-IP) of endogenous proteins would not inform of astrocytic protein-protein associations as SAPAP3, Glt-1, and Ezrin could be expressed in other cell types. Therefore, to validate *in vivo* associations of SAPAP3 with its putative interactors, we first used tagged recombinant proteins expressed in striatal astrocytes for co-IP. We designed three astrocyte-specific AAV constructs and microinjected them into the striatum in different combinations: HA-tagged SAPAP3 (AAV 2/5 *gfaABC$_1$D* 3xHA-SAPAP3), GFP-tagged Glt-1

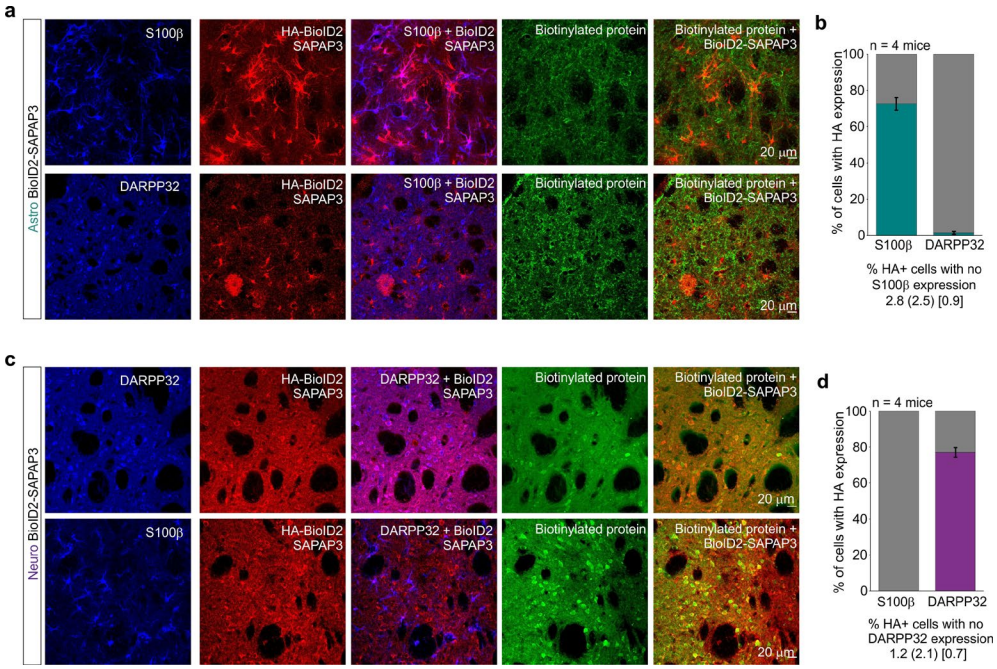

Figure 5.6: Validation of cell-specific SAPAP3-BioID2. **a.** Representative images of immunostained mouse striatum injected with astrocyte-specific BioID2-SAPAP3 and then treated with biotin for 7 days. Top panel shows the immunostaining pattern with S100β as an astrocyte cell marker and bottom panel shows the immunostaining pattern with DARPP32 as a neuronal cell marker. **b.** Bar graphs depicting the percent of S100β positive or DARPP32 positive cells with HA expression in a 40x magnification field of view. Teal portion of the bar graphs show the percent co-localization. (n = 8 fields of view at 40x magnification from 4 mice) **c.** Representative images of immunostained mouse striatum injected with neuron-specific BioID2-SAPAP3 and then treated with biotin for 7 days. Top panel shows the immunostaining pattern with DARPP32 as a neuronal cell marker and bottom panel shows the immunostaining pattern with S100 β as an astrocyte cell marker. **d.** Bar graphs depicting the percent of S100β positive or DARPP32 positive cells with HA expression in a 40x magnification field of view. Purple portion of the bar graphs show the percent co-localization. (n = 8 fields of view at 40x magnification from 4 mice

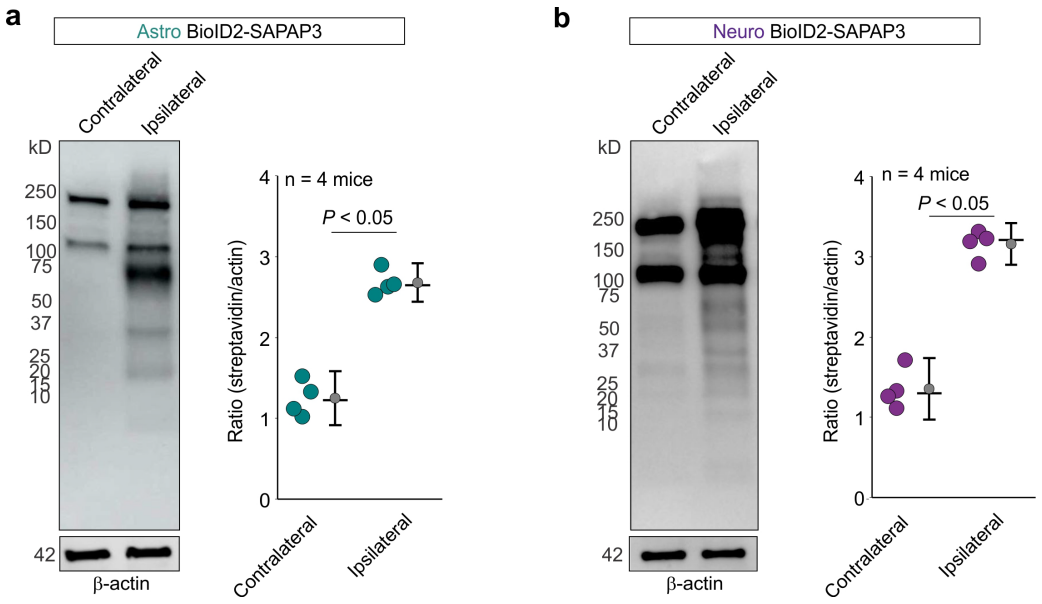

Figure 5.7: Validation of cell-specific SAPAP3-BioID2 with western blots. a. Western blot analysis of brain unilaterally microinjected with astrocyte specific BioID2-SAPAP3. Dark bands at 130 kD and 250 kD show the endogenously biotinylated mitochondrial proteins, Pyruvate carboxylase and acetyl-CoA carboxylase. Graph depicting the streptavidin signal intensity divided by the β-actin signal intensity for each data point. Black horizontal line depicts the mean. (n = 4 mice; Mann-Whitney Test) **b**. Western blot analysis of brain unilaterally microinjected with neuron specific BioID2-SAPAP3. Graph depicting the streptavidin signal intensity divided by the β-actin signal intensity for each data point. Black horizontal line depicts the mean. (n = 4 mice; Mann-Whitney test).

(AAV 2/5 *gfaABC$_1$D* Glt1-GFP), and GFP-tagged Ezrin (AAV 2/5 *gfaABC$_1$D* Ezr-GFP). We injected a set of mice with either HA-SAPAP3 + Glt1-GFP or HA-SAPAP3 + Ezr-GFP and injected another set of mice with each construct separately (**Figure 5.10**). We found that HA-tagged SAPAP3 co-immunoprecipitated with both Glt1-GFP and Ezr-GFP. Similarly, we found that both Glt1-GFP and Ezr-GFP co-immunoprecipitated with HA-tagged SAPAP3. Thus, these results suggest that SAPAP3 does indeed interact with both Glt1 and Ezrin within astrocytes in concordance with our mass spectrometry data.

To explore interactions between endogenous proteins within astrocytes, we used the proximity ligation assay (PLA). In short, the PLA immunoassay uses primary antibodies raised in different species to recognize two proteins of interest. Secondary antibodies against the primary antibodies, called the PLA probes, bind to the primary antibodies. Each of the PLA probes has a short DNA primer that is complimentary to the other PLA probe's primer. If the proteins of interest are associated with each other within a 40 nm range, then, upon the addition of a ligase, a DNA polymerase, and fluorescent oligonucleotides, the polymerase will use the primers to conduct rolling circle DNA synthesis and a resultant fluorescent signal will emerge (Söderberg et al., 2008). To conduct the PLA we used both wild-type mice and SAPAP3 constitutive knockout mice (SAPAP3 KO) as a control for our probes and antibodies. Additionally, the mice were microinjected with astrocyte specific tdTomato AAV (AAV 2/5 *gfaABC$_1$D*-tdTomato) in the striatum for assessment of the PLA within astrocytes and their territories. We found clear PLA puncta signal and thus, associations between endogenous SAPAP3 and Glt-1 within wild-type tdTomato+ astrocyte territories while these signals were significantly absent from tdTomato+ astrocyte territories in SAPAP3 KO mice (**Figure 5.11**). Similarly, we found discrete PLA puncta signal, and thus, associations between endogenous SAPAP3 and Ezrin within wild-type

tdTomato+ astrocyte territories while these signals were significantly absent from tdTomato+ astrocyte territories in SAPAP3 KO mice (**Figure 5.11**).

Therefore, our results indicate that SAPAP3 interacts spatially with both Ezrin and Glt-1 within astrocytes *in vivo.*

Section 5.5: Molecular mechanisms of astrocytic SAPAP3

Together, our proteomic data in conjunction with our validation data indicates that astrocytic SAPAP3 is associated with the plasma membrane via its interactions with the plasma membrane proteome (Chapters 3 and 4) and with its bona-fide interactions with Ezrin and Glt-1. Because of SAPAP3's association with the astrocyte plasma membrane, we next sought to assess whether the plasma membrane proteome of astrocytes was altered in SAPAP3 KO mice relative to wild-type (WT) mice. To do this, we microinjected our astrocyte-specific, plasma membrane targeted BioID2 AAV construct (Astro Lck-BioID2) in the striatum of both SAPAP3 KO and WT mice. To generate a high confidence list of differentially displayed proteins, we also injected a set of SAPAP3 KO and WT mice with Astro Lck-GFP as controls for Astro Lck-BioID2. After mass spectrometry and statistical analyses comparing SAPAP3 KO (Lck-BioID2 vs Lck-GFP) and WT (Lck-BioID2 vs Lck-GFP) groups within each other ($Log_2FC > 1$ and FDR < 0.05 Lck-BioID2 versus Lck-GFP), we compared the astrocyte plasma membrane proteomes of SAPAP3 KO to WT. We found a total of 457 proteins that were altered in the SAPAP3 KO astrocyte plasma membrane proteome when compared to WT. Of these 182 were down, or less associated with the plasma membrane ($Log_2FC < 0.5$ and FDR < 0.05) and 275 were up, or more associated with the plasma membrane ($Log_2FC > 0.5$ and FDR < 0.05). Within these altered proteins, we found the astrocytic SAPAP3 interactors, Ezrin and Glt-1 (**Figure 5.12**). Functional analysis of the 457

altered proteins using Enrichr gene ontology analysis revealed that the most significant dysregulated pathway was “actin cytoskeleton organization” in concordance with the main function of astrocytic SAPAP3 interactors and the SAPAP3-Ezrin interaction (**Figure 5.12**).

SAPAP3 interacts with actin cytoskeleton binding proteins, including Ezrin, in astrocytes. Additionally, upon knock out of SAPAP3, proteins related to actin cytoskeleton organization are dysregulated and the SAPAP3-Ezrin interaction completely disappears. To explore the direct functional interaction of astrocytic SAPAP3 with the actin cytoskeleton, we used LifeAct-GFP (Riedl et al., 2008) as an actin cytoskeleton reporter in both wild-type (WT) and SAPAP3 KO mice. LifeAct is a small 17-amino-acid peptide that binds filamentous actin (F-actin) without disrupting polymerization and depolymerization events. We created an astrocyte-specific AAV construct for LifeAct GFP (AAV 2/5 *gfaABC$_1$D* LifeAct-GFP) and microinjected it into the striatum of both SAPAP3 KO and WT mice at five and a half months of age. After three weeks, we were able to visualize the actin cytoskeleton of individual striatal astrocytes using confocal microscopy. Our results revealed that there was a strong reduction in the actin cytoskeleton within astrocytes from SAPAP3 KO mice. This reduction was the greatest at the peripheral processes of astrocyte territories where the fine processes contact synapses. Additionally, when we measured the total actin area of SAPAP3 KO astrocytes, the entire actin territory of the astrocyte was reduced in comparison to WT (**Figure 5.13**). Analysis of the intensity of the actin signal across a 40 μm area corroborated the LifeAct area measurements because F-actin intensity at the edges of astrocytes were significantly reduced (**Figure 5.13**). We confirmed the smaller territory sizes of astrocytes in SAPAP3 KO mice by assessing the cytosolic tdTomato+ signal within SAPAP3 KO and WT astrocytes. In concordance with the LifeAct data, tdTomato+ astrocytes are smaller in SAPAP3 KO mice when compared to WT (**Figure 5.13)**. In all, our results confirm that astrocytic

SAPAP3 mechanisms directly regulate the actin cytoskeleton, especially at astrocyte fine processes which lie at the most distal edges of a morphologically complex astrocyte.

Section 5.6: Summary

Here, for the first time, we explored the functional differences between a protein that is expressed in two different cell types, astrocytes and neurons, and define different molecular mechanisms for each. Using a cell and subcompartment specific proteomics approach (**Chapters 3 and 4**), we discovered that SAPAP3, a protein previously only associated with the PSD in neurons, is expressed in striatal astrocytes. Validation with multi-omics methods and traditional tissue fluorescence methods provide strong evidence for the presence of SAPAP3 within astrocytes. The data presented here also show that astrocytic SAPAP3 has its own set of unique molecular interactors when compared to neurons, and that these cell-specific interactions give rise to functional mechanisms that relate to actin cytoskeleton organization. We also discovered that SAPAP3 deletion causes astrocytes to become smaller in size and lose their complexity, especially at their fine peripheral processes.

I began this chapter by introducing the concept of proteins working in conjunction with one another to perform a specialized function. In the example of the post-synaptic density, the diverse proteins that orchestrate post-synaptic signaling including the ionotropic receptors, the metabotropic receptors, the adhesion molecules, the scaffolding proteins, and the actin proteins, all work together to integrate and process information so that the post-synaptic neuron can respond and change accordingly. Similarly, as is the case for the PSD, all cells within a circuit must work

Filters	Astro BioID2-SAPAP3	Neuro BioID2-SAPAP3
Total # unique peptides	26,783	
Total # of unique proteins	2,655	
# unique peptides	11,363	15,420
Remove mitochondrial proteins	968	1,459
(vs GFP) log_2ratio >1 and > 3 experiments	216	803
(vs GFP) FDR < 0.05	170	423

Table 5.1: Astrocyte and neuron SAPAP3-BioID2 proteomics analysis filters. List of filters and resultant number of proteins used to generate high confidence lists for both astrocyte and neuron specific SAPAP3-BioID2 proteomes.

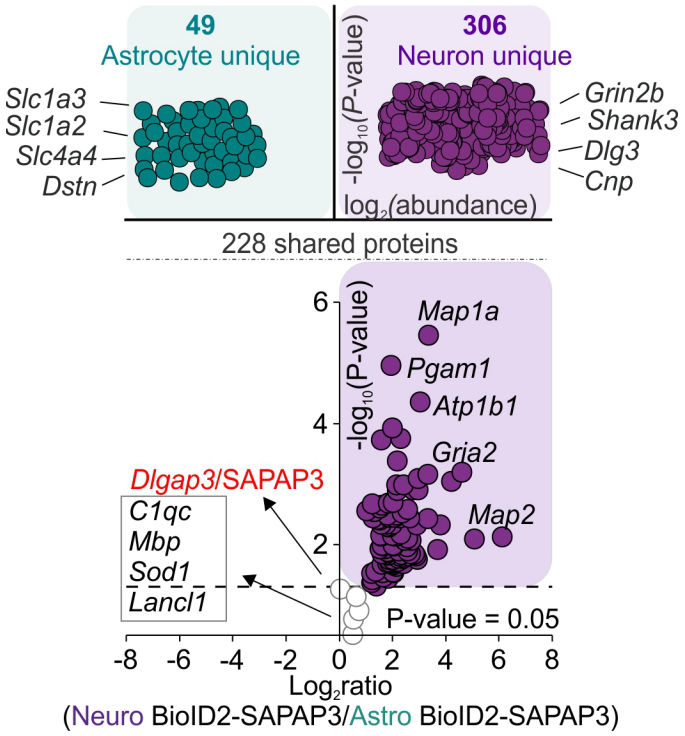

Figure 5.8: Quantitative comparison of cell specific SAPAP3 interactors. Label-free based quantification comparison of significant proteins ($Log_2FC > 1$ and FDR < 0.05 *versus* GFP controls) detected in the Astro BioID2-SAPAP3 and Neuro BioID2-SAPAP3 reveal unique astrocyte and neuron SAPAP3 interactors. Top half of the volcano plot shows 306 unique Neuro BioID2-SAPAP3 proteins and 49 unique Astro BioID2-SAPAP3 proteins when compared to each other. The top four most abundant proteins for each cell type are shown. Lower half of volcano plot shows comparison of 228 proteins that were common in both Astro BioID2-SAPAP3 and Neuro BioID2-SAPAP3. The five highest enriched proteins ($Log_2FC > 2$) for neurons are shown. Proteins that did not pass the enrichment threshold for either cell type are represented in the gray box.

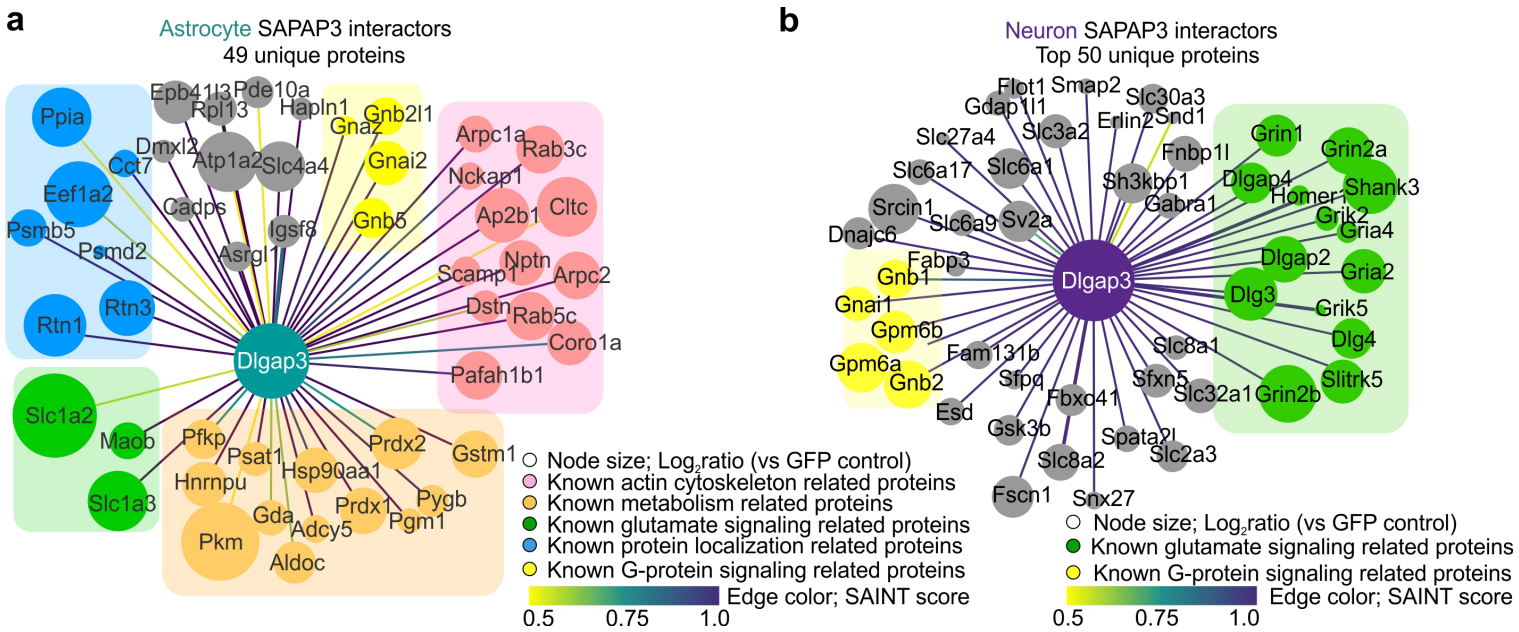

Figure 5.9: SAPAP3 interactome in striatal astrocytes and neurons. a. Predicted interaction map of SAPAP3 interacting astrocyte proteins. Edge color represents SAINT interaction score. Color legends denote PANTHER gene ontology terms. **b.** Predicted interaction map of SAPAP3 interacting neuron proteins. The top 50 by abundance are shown. . Edge color represents SAINT interaction score. Color legends denote PANTHER gene ontology terms.

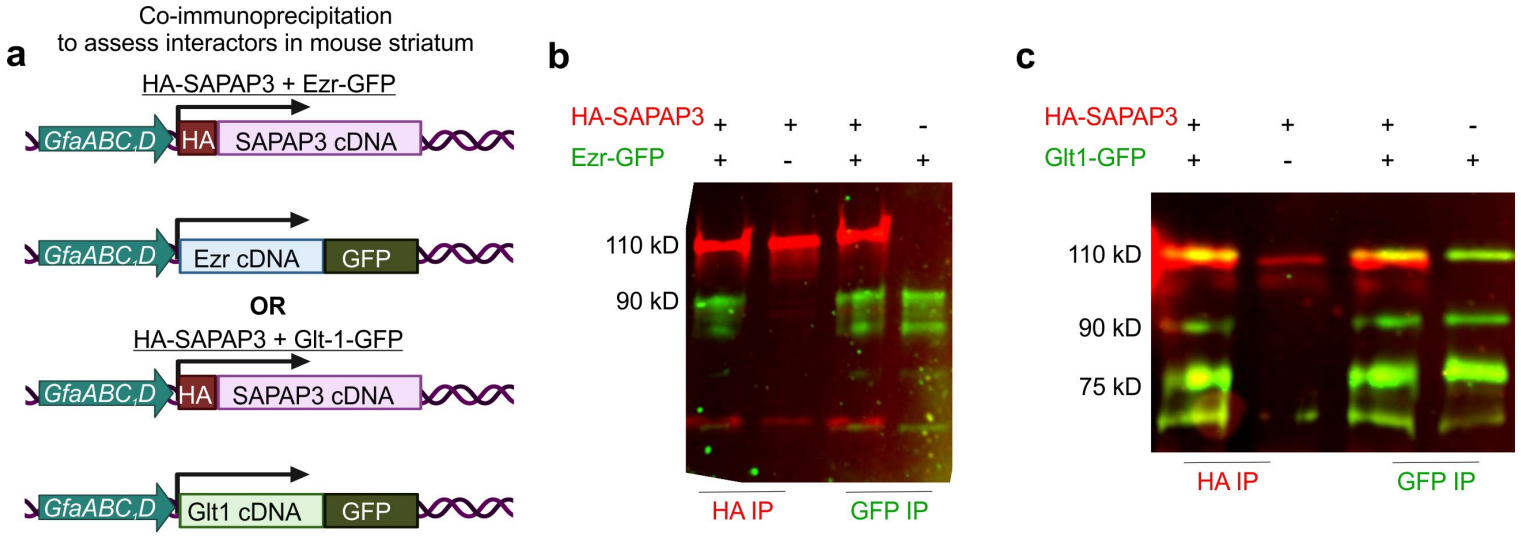

Figure 5.10: Validation of astrocytic SAPAP3 interacting proteins with Co-IP a. Schematic shows astrocyte specific HA tagged SAPAP3, GFP fused Ezrin, and GFP fused Glt-1 used in AAV constructs to assess interactions via co-immunoprecipitation. 16 week old wild type mice were injected in the striatum with one of the following combinations: HA-SAPAP3 + Ezr-GFP, HA-SAPAP3 + Glt-1-GFP, HA-SAPAP3 only, Ezr-GFP only, or Glt1-GFP only. **b and c.** Western blots shows the immunoprecipitation of either HA or GFP after protein complex isolation. The band 110 kD represents the HA-SAPAP3 band, while the 90 kD bands represent Ezrin-GFP (93 kD) or Glt1-GFP (92 kD).

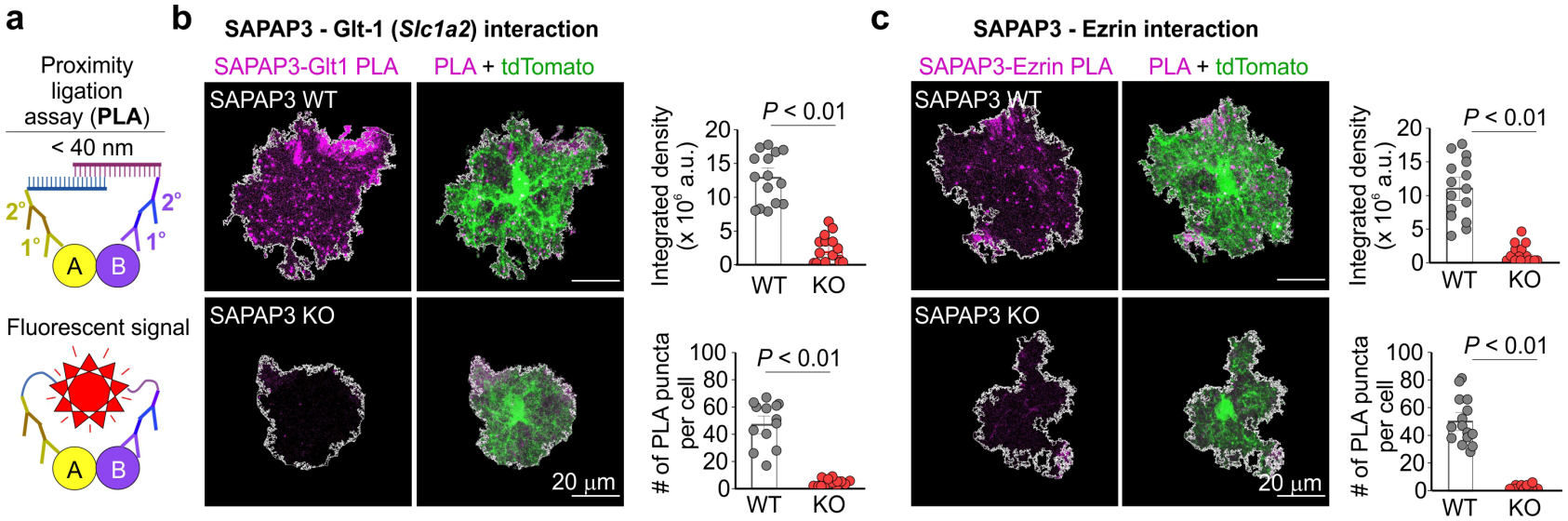

Figure 5.11: Validation of astrocytic SAPAP3 interacting proteins with PLA a. Cartoon depicting the principle of the proximity ligase assay. **b.** Images of PLA puncta for SAPAP3 and Glt-1 within tdTomato+ astrocytes from wild-type and SAPAP3 KO mice. Graphs show the integrated density and number of PLA puncta per cell. (n = 15 tdTomato+ astrocytes from 3 mice per group; two-tailed Mann-Whitney test). **c.** As in d, but for SAPAP3 and Ezrin proteins within tdTomato+ astrocytes from wild-type and SAPAP3 KO mice. (n = 15 tdTomato+ astrocytes from 3 mice per group; two-tailed Mann-Whitney test).

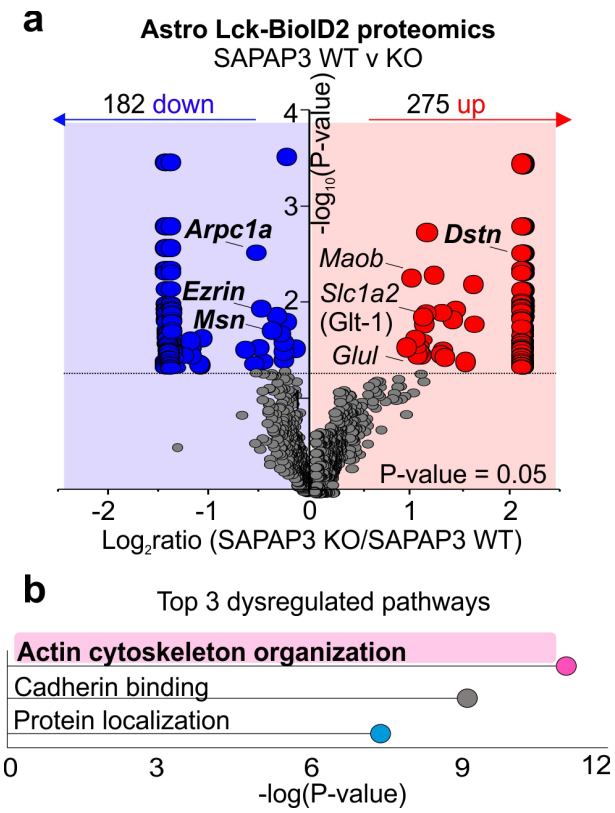

Figure 5.12: Comparison of astrocyte plasma membrane proteomes in wild-type and SAPAP3 KO mice. a. Volcano plot showing differentially displayed astrocyte plasma membrane proteins in SAPAP3 KO mice when compared to wild-type. **b.** Graph shows most significant Enrichr gene ontology terms for "molecular function" from the 457 differentially displayed astrocyte proteins.

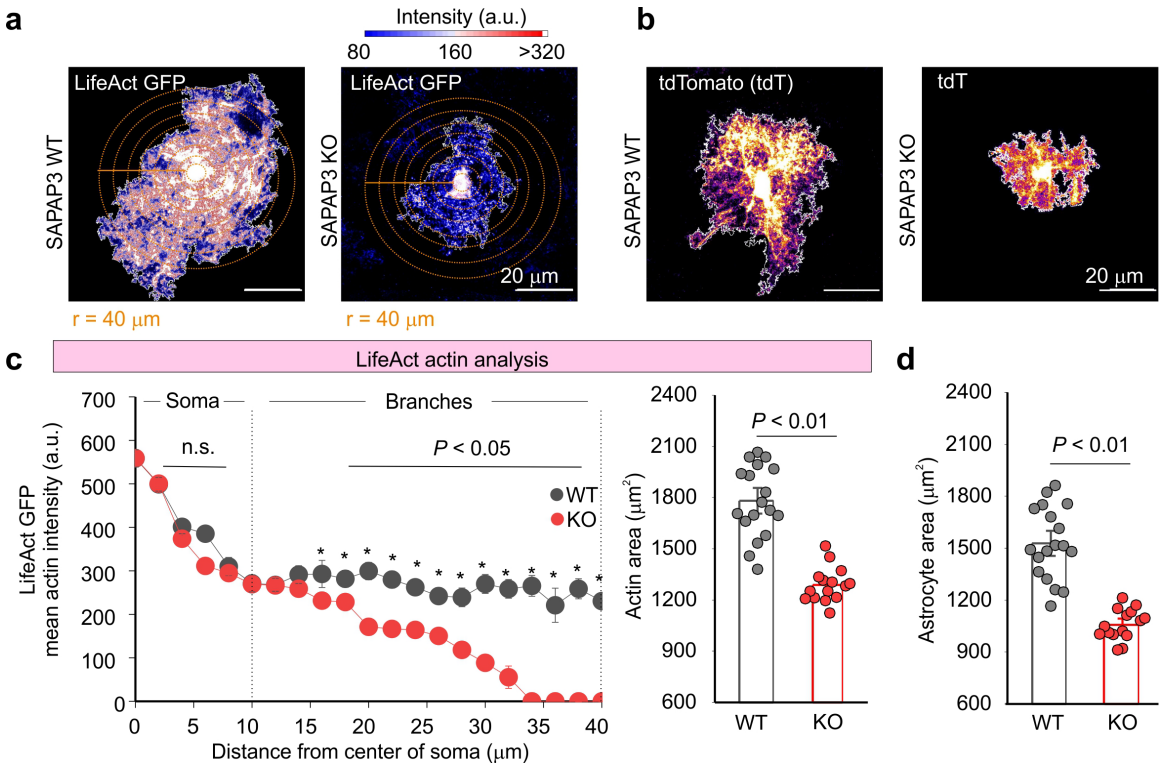

Figure 5.13: Astrocyte morphology is disrupted in SAPAP3 KO mice. a. Representative images showing LifeAct GFP expression in astrocytes of wild-type and SAPAP3 KO mice. Concentric circles represent where intensity was measured as a function of distance from the soma center. **b.** Representative images showing wild-type and SAPAP3 KO tdTomato+ astrocytes. **c.** Line graph depicting LifeAct GFP intensity as a function of distance from the center of the astrocyte soma. Points represent mean intensity from 15-18 cells per group from 4 mice. (Two-way repeated measures ANOVA with Bonferroni post-hoc test; P = 0.012 at 20-40 µm). Bar graphs shows the total LifeAct GFP area in wild-type and SAPAP3 KO mice. **d.** Bar graph shows the tdTomato+ astrocyte area. (n = 15 WT and 18 SAPAP3 KO tdTomato+ astrocytes from 4 mice per group. The mean and SEM are shown; Two-tailed unpaired t-test with Welch correction).

together for an action to be produced at the organismal level. Our SAPAP3 molecular experiments in both astrocytes and neurons shown here provide a clue on how cells may work together to produce an intended output. SAPAP3 expressed in both astrocytes and neurons has conserved functions in both cell types: glutamatergic signaling and G-protein signaling. However, the proteins composing these conserved functions are completely different in both cell types. Therefore, striatal SAPAP3 may work in different cell types to regulate similar pathways in order keep striatal function working toward its goal: regulating thalamic input into the cortex. However, how can different cell types collaborate with neurons at the circuit-level to keep working at this goal?

While we only describe molecular mechanisms here, we must understand what occurs at the organismal level in order to dissect cell-specific contributions of astrocytes and neurons that keep the striatal circuit functioning as intended. To probe this, we assessed the behavior of SAPAP3 KO mice in relation to wild-type, as SAPAP3 KO mice are a model of obsessive-compulsive disorder (Welch et al., 2007). Thus, we seek to understand whether the molecular mechanisms presented here for astrocytes and neurons have any relevance for striatum-related behaviors that have gone awry.

CHAPTER 6: ASTROCYTE AND NEURON SPECIFIC CONTRIBUTIONS IN A MOUSE MODEL OF OBSESSIVE-COMPULSIVE DISORDER

Section 6.1: Introduction

SAPAP3, a cytosolic protein expressed in both striatal astrocytes and medium spiny neurons (MSNs), is one of 63 genes associated with obsessive-compulsive disorder (OCD) and repetitive behaviors in humans (Bienvenu et al., 2009; Zuchner et al., 2009; Yu et al., 2010; Stein et al., 2019). Obsessive-compulsive disorder (OCD) is a debilitating psychiatric disorder with a lifetime prevalence that affects about 2% of the population and primarily arises during adolescence or early adulthood (nimh.nih.gov). OCD is primarily characterized by obsessive thoughts and/or compulsions. Additionally, these obsessive thoughts are often associated with anxiety. Cognitive-behavioral theories emphasize that obsessions can often lead to an increase in anxiety and thus, compulsions are a response to the obsessive thoughts (Stein et al., 2019). These defects in cognitive processing in patients with OCD are mediated by alterations in specific neural circuitry. Advances in both structural and functional brain imaging have given rise to important models related to OCD neurocircuitry. Such data have attributed OCD to arise in parallel cortico-striato-thalamo-cortical (CSTC) circuits that are involved in a range of sensorimotor, cognitive, and motivational processes. Alterations in several regions that comprise the CSTC circuitry have been demonstrated in OCD patients when compared with healthy, control individuals (Saxena and Rauch, 2000; Rotge et al., 2008; Kwon et al., 2009).

Therefore, to study OCD, several mouse models which show defects in CSTC circuitry have arisen. One of these models, is the constitutive SAPAP3 KO mouse model of OCD (Welch et al., 2007). SAPAP3 KO mice are characterized by repetitive and excessive self-grooming that leads to facial lesions accompanied by severe anxiety. The repetitive and excessive self-grooming arises

at about 5 to 6 months (~P180) of age and can be ameliorated with Fluoxetine, a selective serotonin reuptake inhibitor (SSRI) and first line of treatment for human OCD (Welch et al., 2007; Burguiere et al., 2013). Additionally, as mentioned previously, SAPAP3 KO mice exhibit alterations in CSTC circuitry, especially at corito-striatal synapses (Burguiere et al., 2013; Wan et al., 2014). Thus, the SAPAP3 KO mouse model presents itself as an ideal resource by which to study OCD-like behaviors in mice given its robust phenotype and the genetic associations of SAPAP3 to human OCD.

Because we have discovered the molecular mechanisms by which SAPAP3 functions in astrocytes and compared these mechanisms to those in neurons (**Chapter 5**), we wanted to assess whether these cell-specific mechanisms were relevant to OCD-related behaviors in SAPAP3 KO mice. Additionally, we sought to explore whether these cell-specific mechanisms were of any consequence to the function of the CTSC circuit. To do this, we genetically targeted striatal astrocytes and MSNs to assess cell-specific contributions of these cell types in the SAPAP3 KO mouse model of OCD.

Section 6.2: SAPAP3 KO mice exhibit OCD-like behaviors

To begin exploring cell-type specific contributions of astrocytes and neurons in the SAPAP3 KO mouse model of OCD, we began by reproducing OCD-like behaviors in both WT and SAPAP3 KO mice at 6 months of age. We assessed repetitive behaviors and self-grooming by measuring the area of facial lesions, the number of facial lesions, the number of self-grooming bouts, and the total time spent grooming. We explored anxiety behaviors by measuring time spent in the center of an open field arena (OFA), average speed in the center of an open field arena (OFA), and time spent exploring open arms in an elevated plus maze (EPM). We measured ambulation and

locomotor activity as total distance traveled and total average speed in an OFA. As expected, 6 month old SAPAP3 KO mice showed a significant increase in both number of lesions and area of lesions when compared to WT mice which did not show any lesions (**Figure 6.1**). Additionally, SAPAP3 KO mice spent about 50% of the time engaging in grooming behaviors including paw licking, face stroking, body fur licking, body scratching with hind paws, tail licking, and genital cleaning. This was also reflected by the increased number of grooming bouts (**Figure 6.1**). When we measured anxiety behaviors we found that SAPAP3 KO mice spent significantly less time in the center of the OFA, traveled faster in the center of the OFA, and spent less time exploring the open arms of the EPM when compared to their WT littermates, thus highlighting the high anxiety behaviors present in SAPAP3 KO mice (**Figure 6.2**). Assessment of total activity measured by locomotor measures revealed that SAPAP3 KO mice ambulate slower and spend less time ambulating when compared to their WT littermates (**Figure 6.2**). Thus, we were able to reproduce SAPAP3 OCD-related behaviors to conduct downstream manipulations.

In order to benchmark any subsequent genetic manipulations to Fluoxetine, which ameliorates OCD-like behaviors in SAPAP3 KO mice and is a first line of treatment in human OCD, we reproduced the effects of fluoxetine in SAPAP3 KO mice. We intraperitoneally injected Fluoxetine (10 mg/kg) or saline as a control for 7 days to 6 month old SAPAP3 KO or WT mice. In this way we had four experimental groups: WT + saline, WT + fluoxetine, KO + saline, and KO + fluoxetine. We assessed the same self-grooming, anxiety, and locomotion behaviors that we used for the SAPAP3 KO mice and found that 7 day fluoxetine treatments decreased the number of grooming bouts and grooming time in KO + Fluoxetine mice when compared to KO + saline mice. However, because the 7 day treatment was acute, we did not observe any changes to the area or number of lesions although they did seem to be scarring (**Figure 6.3**). Fluoxetine did not have any

effect on any of our measured grooming parameters in WT mice. When we assessed anxiety behaviors, 7 day Fluoxetine treatment caused SAPAP3 KO mice to increase their time spent in the center of the OFA, decrease their speed in the center of the OFA, and increase the time they spent exploring the open arms in the EPM when compared to KO + saline mice. This indicates that fluoxetine is able to rescue the anxiety behaviors in SAPAP3 KO mice (**Figure 6.4**). As with grooming, fluoxetine did not have any effect on any of our measured anxiety parameters in WT mice. When we measured the locomotor activity of Fluoxetine treated mice, both total distance traveled and total average speed were significantly increased in KO + Fluoxetine mice when compared to KO + saline mice, indicating that Fluoxetine also improves ambulation with no effects found in the WT mice (**Figure 6.4**). Overall, our data show that we are able to reproduce both the self-grooming and anxiety behaviors that characterize the SAPAP3 KO model mice at 6 months of age, and that we could reproduce the effects of Fluoxetine in these mice. In this way, we now have benchmark data by which we can assess the effects of genetic manipulations or treatments.

Section 6.3: Astrocyte and neuron-specific genetic rescue of SAPAP3

As mentioned previously, a highlight of SAPAP3 KO mice is their altered cortico-striato-thalamo-cortical (CSTC) circuitry, with the most significant changes occurring in the striatum, more specifically at cortico-striatal synapses (Burguiere et al., 2013; Wan et al., 2014). Given that SAPAP3 is highly enriched within the striatum (Welch et al., 2004) and the striatum is a critical

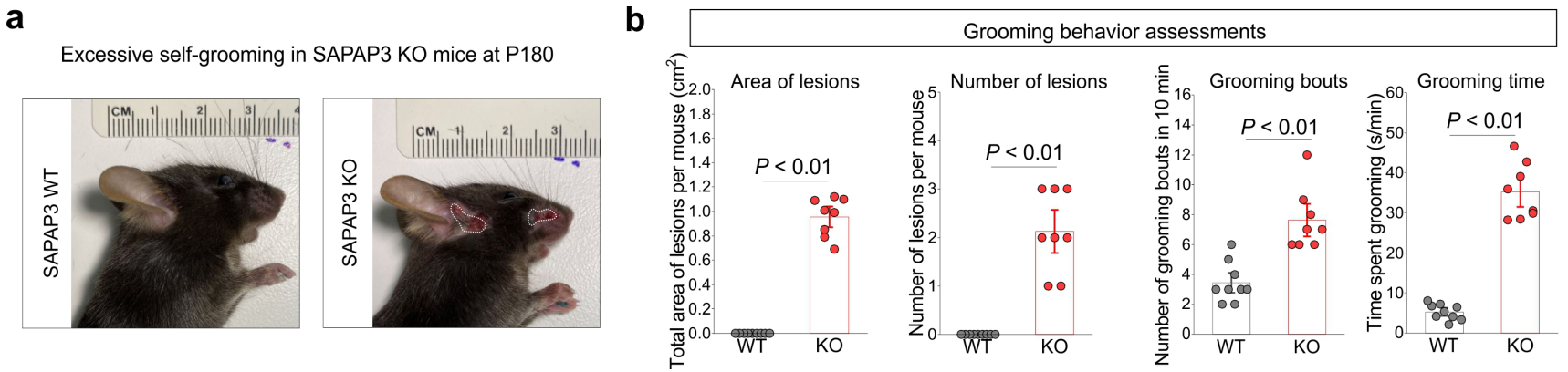

Figure 6.1: SAPAP3 KO mice exhibit excessive self-grooming. a. Representative photographs of WT and SAPAP3 KO mice at P180. White outline shows the border of each lesion. **b**. Bar graphs show assessment of grooming behavior in WT and SAPAP3 KO by four different metrics: area of open lesions, number of lesions, grooming bouts, and time spent grooming. N = 9 for the WT mice and N = 8 for the SAPAP3 KO mice. (Area of open lesions: WT v KO – One sample t-test. Number of lesions per mouse: WT v KO – One sample t-test. Number of grooming bouts: WT v KO - Mann-Whitney test. Time spent grooming: WT v KO – Unpaired t-test with Welch correction.)

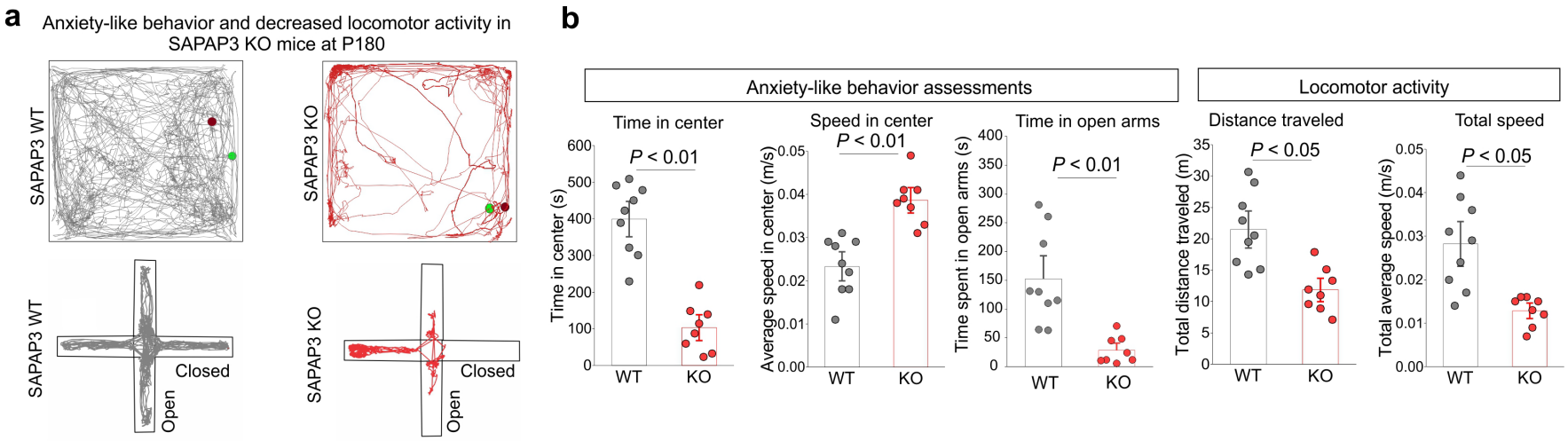

Figure 6.2: SAPAP3 KO mice exhibit high anxiety. a. Representative traces of 20 minute open field recordings for WT and SAPAP3 KO mice at P180. Green dot depicts the location of the mouse at the start of the recording, while the red dot depicts the location of the mouse at the end of the recording. Bottom traces are of 10 minute elevated plus maze recordings. **b**. Bar graphs show locomotor activity and anxiety-like behavior in WT and SAPAP3 KO mice by five different metrics: total distance traveled, total average speed, time in center of the open field, speed in center of the open field, and time spent in the open arms of the elevated plus maze. n = 9 for the SAPAP3 WT mice and n = 8 for the SAPAP3 KO mice. (time in center: WT v KO, two-tailed unpaired t-test with Welch correction, $P < 0.001$; speed in center: WT v KO, two-tailed unpaired t-test with Welch correction, $P = 0.001$; time in open arms: WT v KO, two-tailed unpaired t-test with Welch correction, $P = 0.002$; distance traveled: WT v KO, two-tailed unpaired t-test with Welch correction, $P = 0.001$; average speed: WT v KO, two-tailed unpaired t-test with Welch correction, $P = 0.002$)

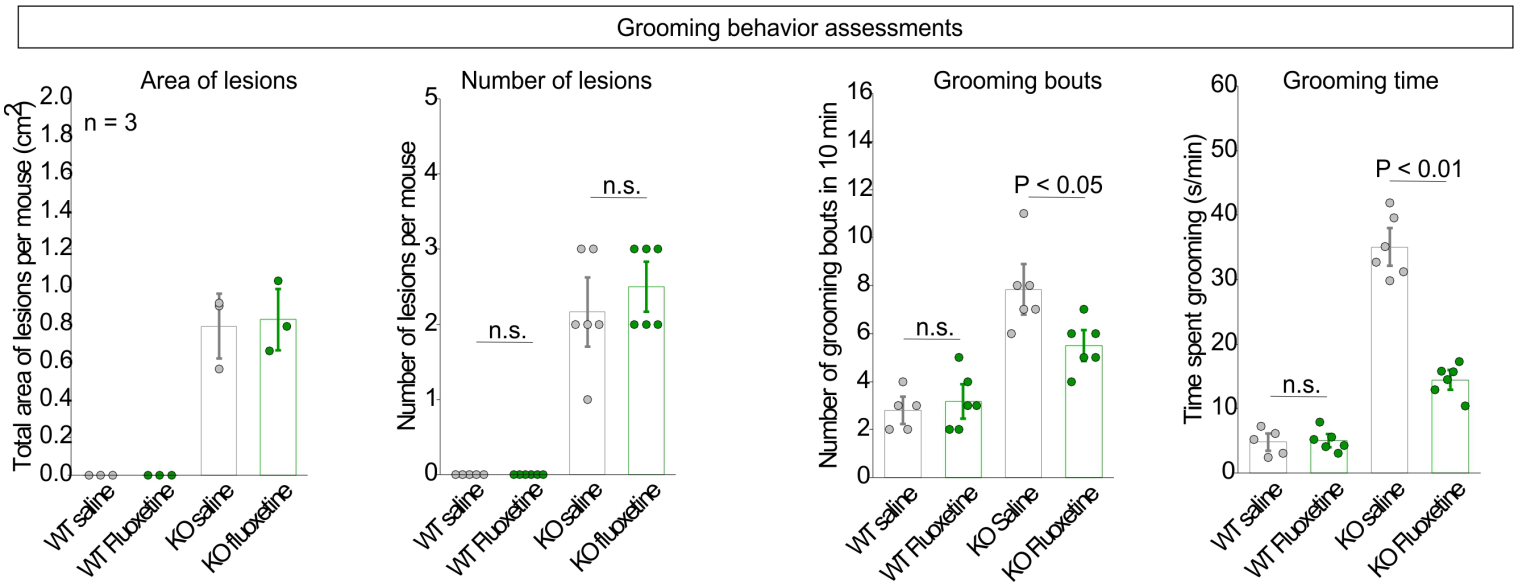

Figure 6.3: Fluoxetine ameliorates excessive self-grooming. Bar graphs show assessment of grooming behavior in SAPAP3 WT and SAPAP3 KO treated with either saline or fluoxetine (10 mg/kg) for 7 days by four different metrics: area of open lesions, number of lesions, grooming bouts, and time spent grooming. n = 5-6 mice for each group. The mean and SEM are shown (number of lesions per mouse: Kruskal Wallis test; number of grooming bouts: One way ANOVA with Bonferroni correction; time spent grooming: One way ANOVA with Bonferroni correction)

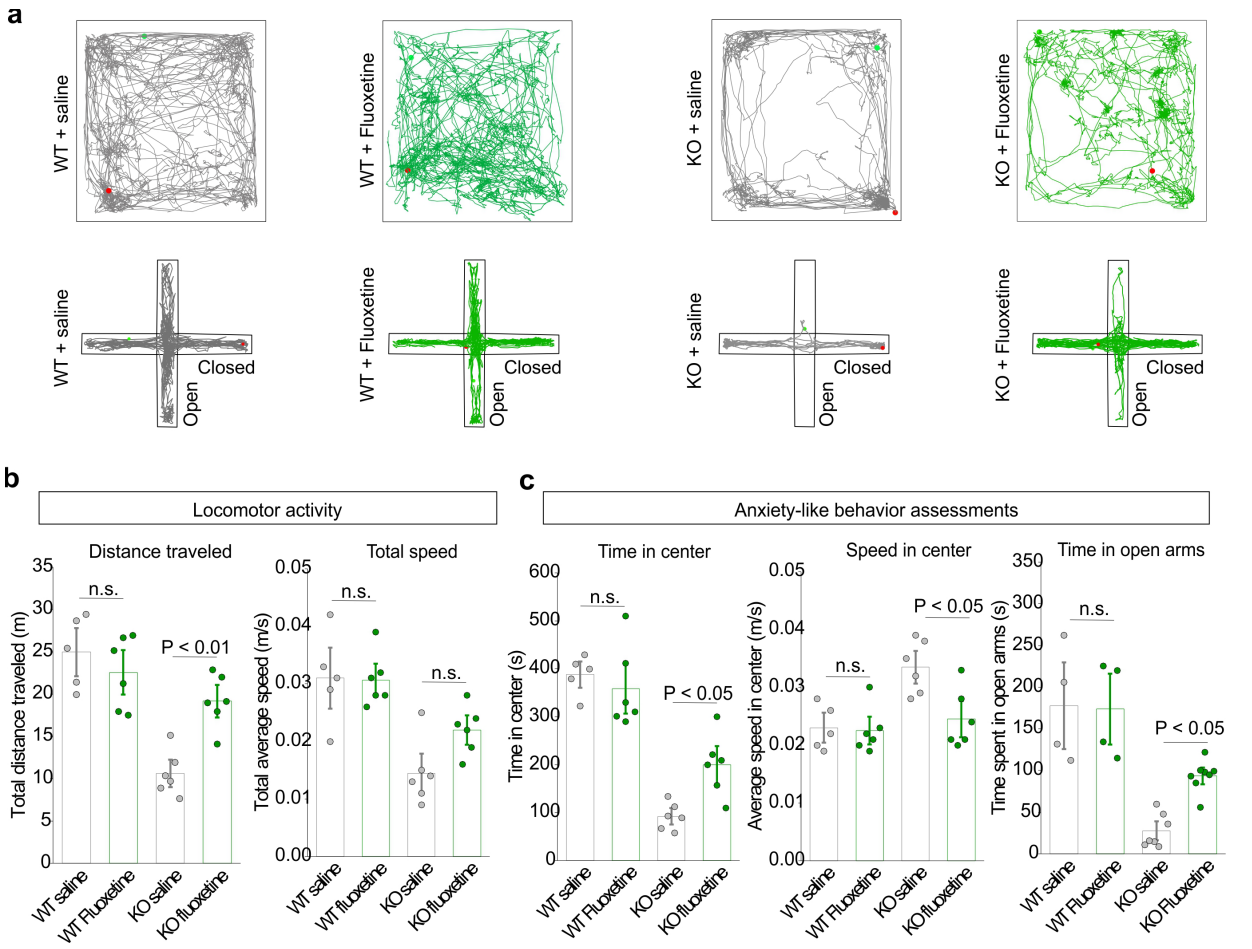

Figure 6.4: Fluoxetine ameliorates anxiety. a. Representative traces of the open field arena and elevated plus maze of SAPAP3 WT and SAPAP3 KO mice treated with either saline or fluoxetine. **b.** Bar graphs show assessment of locomotor activity in SAPAP3 WT and SAPAP3 KO treated with either saline or fluoxetine for 7 days by two different metrics: total distance traveled and total average speed. n = 5 mice WT + saline, n = 6 mice WT + fluoxetine, n = 6 mice KO + saline, n = 6 mice KO + fluoxetine. The mean and SEM are shown (total distance traveled: one way ANOVA with Bonferroni correction; total average speed: one way ANOVA with Bonferroni correction). **c.** Bar graphs show assessment of anxiety-like behaviors in SAPAP3 WT and SAPAP3 KO treated with either saline or fluoxetine for 7 days by three different metrics: time in center, speed in center, and time spent in open arms. n = 4 mice WT + saline, n = 4 mice WT + fluoxetine, n = 7 mice KO + saline and n = 8 mice KO + fluoxetine. The mean and SEM are shown (time in center: One way ANOVA with Bonferroni correction; speed in center: one way ANOVA with Bonferroni correction; time spent in open arms: Kruskal-Wallis test).

component of basal ganglia circuitry that gives rise to repetitive behaviors such as those arising in SAPAP3 KO mice (Graybiel, 2008; Burguiere et al., 2015), direct manipulations within the striatum could provide beneficial effects. Indeed, when SAPAP3 KO mice receive direct expression of SAPAP3 only in striatum, both anxiety and repetitive behaviors are rescued (Welch et al., 2007). In these studies, a GFP-SAPAP3 fusion protein or a GFP only control was packaged into a lentiviral vector. These vectors were then locally microinjected bilaterally and broadly into the striatum at a young age (P7) and behaviors were assessed at 6 months of age. Indeed, striatal expression of GFP-SAPAP3 rescues both excessive self-grooming and anxiety in SAPAP3 KO mice when compared to GFP only controls (Welch et al., 2007). While this seminal study demonstrated that the striatum is the major driver of altered behaviors in SAPAP3 KO mice, the vectors used were not cell-selective, thus the contributions of each cell type that expresses SAPAP3 to these SAPAP3 KO related behaviors is unknown.

Because we discovered that striatal astrocytes, in addition to neurons, express SAPAP3, we sought to assess the contribution of both cell types to repetitive and anxiety behaviors in SAPAP3 KO mice. To do this we designed four cell-selective AAV vectors: an astrocyte specific GFP-SAPAP3 fusion protein (AAV 2/5 *gfaABC$_1$D* GFP-SAPAP3) hereafter referred to as Astro SAPAP3; an astrocyte specific GFP control (AAV 2/5 *gfaABC$_1$D* GFP) hereafter referred to as Astro GFP; a neuron specific GFP-SAPAP3 fusion protein (AAV 1 *hSyn1* GFP-SAPAP3) hereafter referred to as Neuro SAPAP3; and a neuron specific GFP control (AAV 1 *hSyn1* GFP). We first tested if these vectors could be broadly expressed in the striatum and whether they were cell-selective. We microinjected Astro SAPAP3 and Neuro SAPAP3 into the striatum of WT mice at P28 at three different sites: ventral striatum, medial striatum, and dorsal striatum. One month later we assessed expression with immunohistochemistry. Both Astro SAPAP3 and Neuro

SAPAP3 were expressed broadly in the striatum (**Figure 6.5**). Additionally, both constructs were cell selective as Astro SAPAP3 was expressed in ~80% of S100β+ astrocytes within the striatum and in ~0% of DARPP32+ neurons. Conversely, Neuro SAPAP3 was expressed in ~70% of DARPP32+ neurons and in 0% of S100β+ astrocytes, indicating that both Neuro SAPAP3 and Astro SAPAP3 are well expressed and cell-selective in the striatum (**Figure 6.5**).

Validation of our cell-selective SAPAP3 constructs prompted us to then explore the cell-specific contributions of astrocytes and neurons to OCD-like behaviors in mice. To do this, we microinjected either cell-selective SAPAP3 or cell-selective control GFP broadly into the striatum of SAPAP3 KO mice at P21-P28 days of age. We then allowed expression for five months or until the GFP control mice began to groom excessively (~P180 or 6 months of age). We then assessed repetitive behaviors and self-grooming by measuring the area of facial lesions, the number of facial lesions, the number of self-grooming bouts, and the total time spent grooming. We explored anxiety behaviors by measuring time spent in the center of an open field arena (OFA), average speed in the center of an open field arena (OFA), and time spent exploring open arms in an elevated plus maze (EPM). We measured ambulation and locomotor activity as total distance traveled and total average speed in an OFA.

Our experiments revealed that at 6 months of age, both Astro SAPAP3 and Neuro SAPAP3 genetic interventions significantly recovered excessive grooming behaviors in SAPAP3 KO mice including: reduced the total area of lesions per mouse, reduced the total number of lesions per mouse, and reduced the total time spent grooming when compared to KO mice microinjected with GFP only. The total number of grooming bouts was significantly reduced with the Astro SAPAP3 intervention, and trended as reduced (no significance) with the Neuro SAPAP3 intervention when

compared with their respective GFP controls (**Figure 6.6**). Thus, this indicates that both astrocytes and neurons play a role in SAPAP3 mediated repetitive behaviors that arise in the striatum.

We next assessed anxiety behaviors in our genetically treated SAPAP3 KO mice. Interestingly, SAPAP3 KO mice microinjected with Neuro SAPAP3 showed increased time in center in the OFA, decreased average speed in the center of the OFA, and increased time spent exploring the open arms of the EPM when compared to mice microinjected with control Neuro GFP. However, SAPAP3 KO mice microinjected with Astro SAPAP3 showed no change in any of the anxiety measurements or behaviors described above when compared to their respective litter-mates microinjected with control Astro GFP, thus suggesting a significant effect of neuronal genetic intervention on the anxiety-like behavior displayed by SAPAP3 KO mice (**Figure 6.7**).

To assess whether this difference of anxiety behaviors in Astro SAPAP3 and Neuro SAPAP3 treated mice was caused by differences in locomotion, we measured the total distance traveled and average speed of the experimental mice. We found that both Astro SAPAP3 and Neuro SAPAP3 significantly increased total distance traveled in the OFA, and increased the total average speed in SAPAP3 KO mice compared to GFP controls and to the same effect. Therefore, this indicates that locomotion is independent from the differences in anxiety behaviors that arise from the cell-specific genetic intervention and further suggest that both astrocyte and neuron SAPAP3 have a direct effect on striatum mediated behaviors such as repetition and locomotion (**Figure 6.7**).

Using our benchmark Fluoxetine data described previously, we then compared our cell-specific genetic interventions compared to Fluoxetine intervention. By using the Fluoxetine rescue as a baseline metric, we found that both Astro SAPAP3 and Neuro SAPAP3 resulted in beneficial effects comparable to fluoxetine for self-grooming with all three interventions producing a 70% recovery (**Figure 6.8**). As described previously, only Neuro SAPAP3 had a comparable effect to

fluoxetine for anxiety with both interventions producing an 80% recovery in the OFA evaluations. These results show that cell-specific genetic rescue of SAPAP3 in either astrocytes or neurons provide an avenue to explore therapeutic strategies that target both cell types in OCD.

In all, our data reveal that in the context of SAPAP3 KO, both striatal astrocytes and neurons have cell-specific contributions that drive OCD-like behaviors including excessive self-grooming and anxiety. More specifically, both astrocytes and neurons contribute to repetitive behaviors defined by excessive self-grooming and development of lesions, as both these readouts are ameliorated with genetic rescue in the striatum via Astro SAPAP3 and Neuro SAPAP3. However, only neuronal intervention with Neuro SAPAP3 benefits anxiety in the OFA and EPM indicating that the manner by which astrocytes and neurons contribute to different behaviors is separable.

Section 6.4: Effects of cell-specific rescue on striatal astrocytes

Given that a cell-specific genetic intervention was sufficient to improve OCD-like behaviors in SAPAP3 KO mice and that astrocyte SAPAP3 molecular mechanisms are disrupted in SAPAP3 KO mice (**Chapter 5**), we assessed the effect of these interventions on striatal astrocytes by assessing their morphology and molecular interactors.

We microinjected SAPAP3 KO mice in the striatum with either of the four SAPAP3 AAV constructs (Astro SAPAP3, Astro GFP, Neuro SAPAP3, Neuro GFP) in combination with astrocyte specific tdTomato (*gfaABC$_1$D* tdTomato) at P21-P28 days of age. We then assessed the

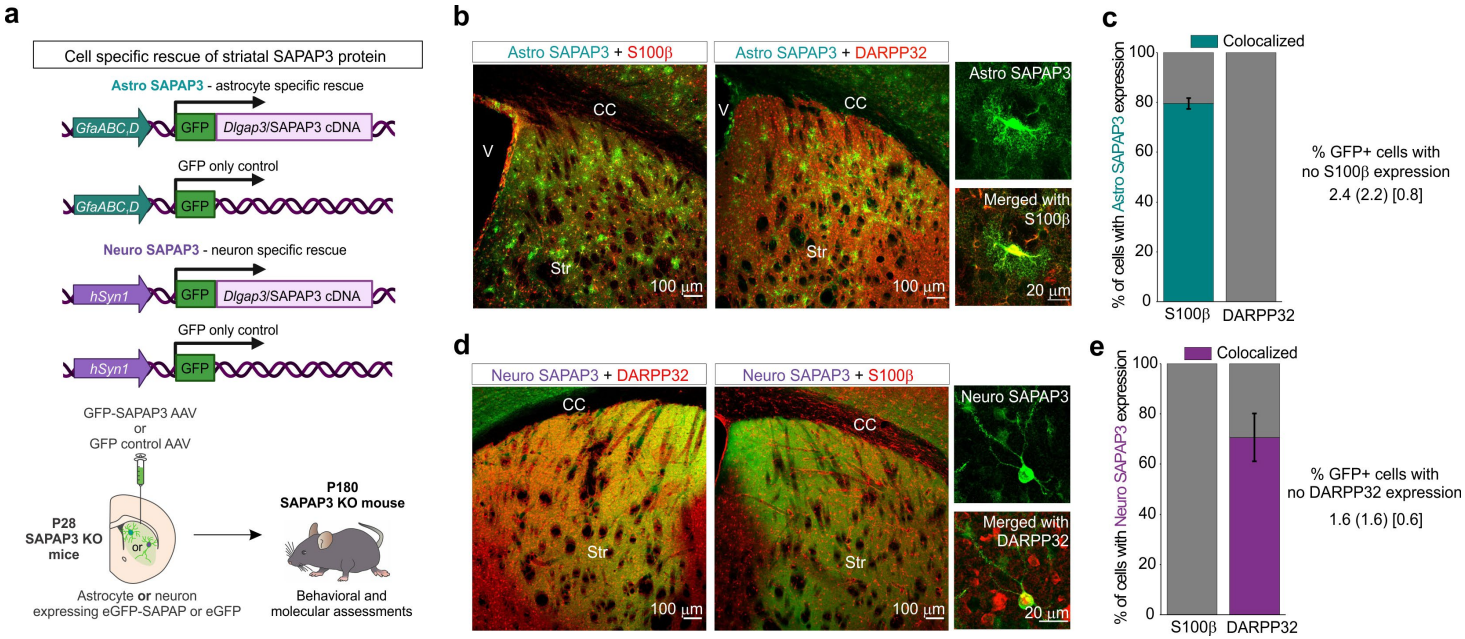

Figure 6.5: Strategy for cell-specific rescue of SAPAP3 a. Cartoon depicting the genetic strategy and experimental plan for cell-specific rescue of SAPAP3 in the striatum of SAPAP3 KO mice. **b.** Representative images of immunostained mouse striatum injected with astrocyte-specific GFP-SAPAP3 (Astro SAPAP3). Left panel shows the immunostaining pattern with S100β as an astrocyte cell marker and right panel shows the immunostaining pattern with DARPP32 as a neuron cell marker. **c.** Bar graphs depicting the percent of S100β positive or NeuN positive cells with HA expression in a 20x magnification field of view. Teal portion of the bar graphs show the percent co-localization. (n = 8 fields of view at 20x magnification from 4 mice). **d.** Representative images of immunostained mouse striatum injected with neuron-specific GFP-SAPAP3 (Neuro SAPAP3). Left panel shows the immunostaining pattern with DARPP32 as a neuron cell marker and right panel shows the immunostaining pattern with S100β as an astrocyte cell marker. **e.** Bar graphs depicting the percent of S100β positive or NeuN positive cells with HA expression in a 20x magnification field of view. Purple portion of the bar graphs show the percent co-localization. (n = 8 fields of view at 20x magnification from 4 mice)

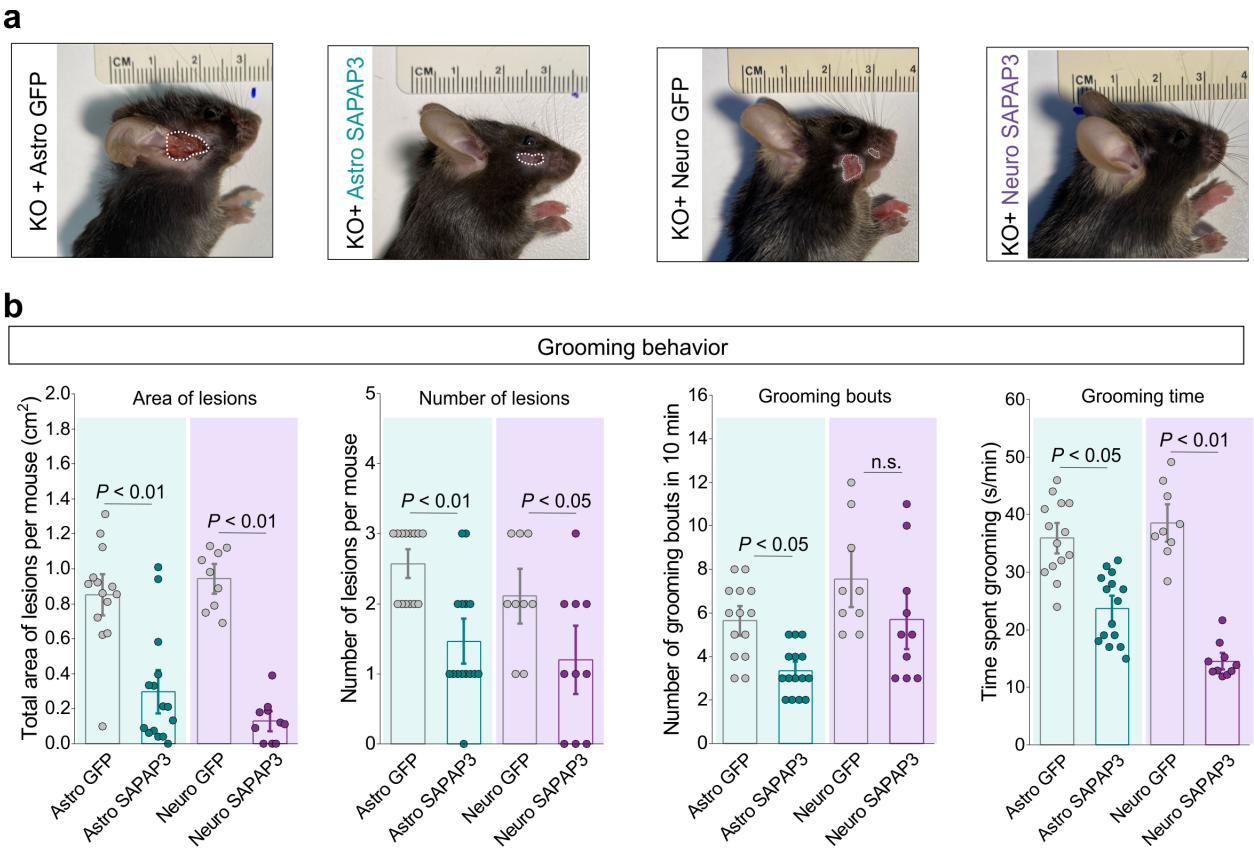

Figure 6.6: Effect of cell-specific genetic rescue on excessive self-grooming. a. Photographs show SAPAP3 KO mice injected with Astro GFP, Astro SAPAP3, Neuro GFP, or Neuro SAPAP3 .White outline shows lesion border. **b**. Grooming behaviors assessed in SAPAP3 KO mice treated with cell-specific GFP or SAPAP3. (All behaviors: n = 9-12 mice per group; Kruskal-Wallis Test with Dunn post-hoc test)

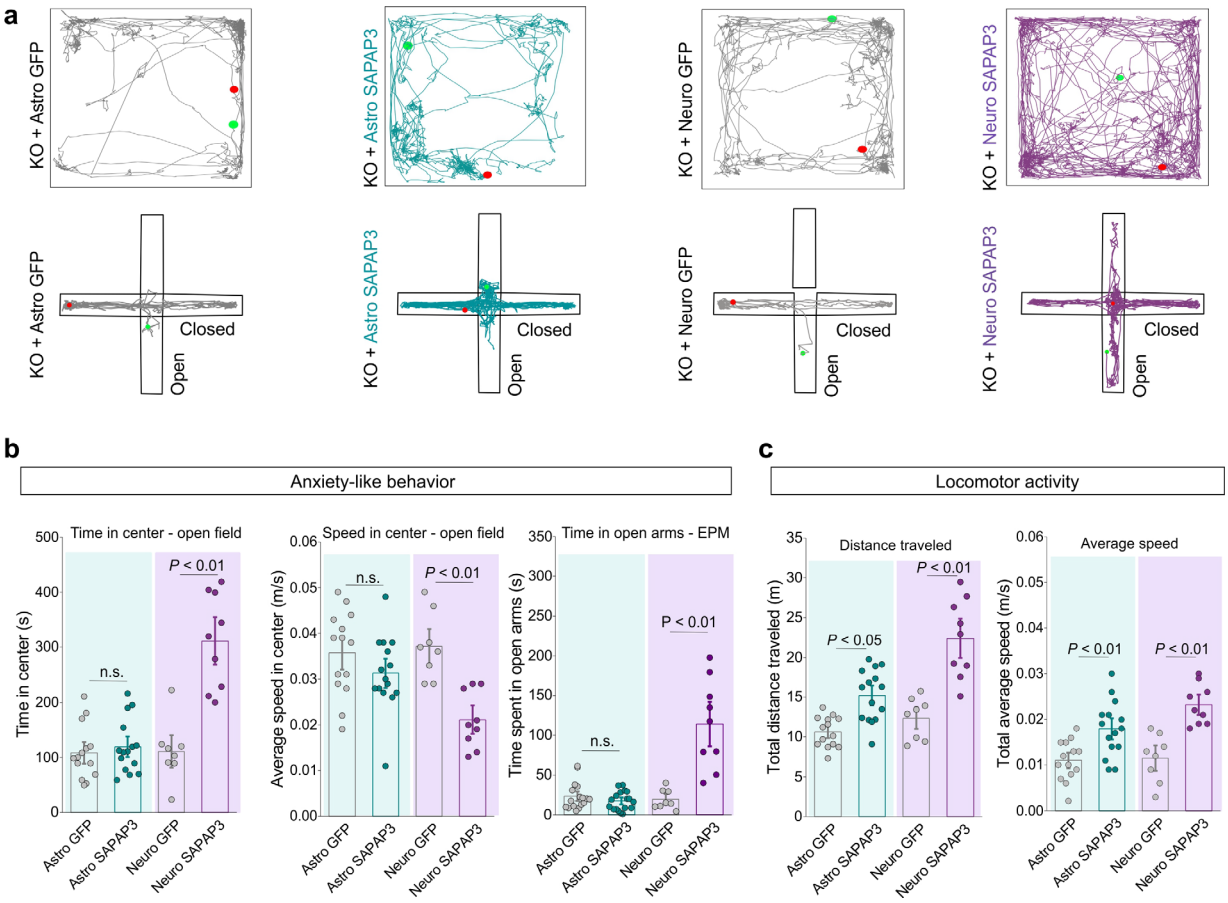

Figure 6.7: Effect of cell-specific genetic rescue on anxiety. a. Representative traces of open field recordings and elevated plus maze recordings of SAPAP3 KO mice injected with Astro GFP, Astro SAPAP3, Neuro GFP, or Neuro SAPAP3. Green dot shows the starting location of the mouse, while the red dot represents the final location. b. Anxiety-like behaviors measured in SAPAP3 KO mice treated with either cell-specific GFP or SAPAP3. (All behaviors: n = 9-12 mice per group; Kruskal-Wallis Test with Dunn post-hoc test). c. Locomotor activity assessed in SAPAP3 KO mice treated with either cell-specific GFP or cell-specific SAPAP3. (All behaviors: n = 9-12 mice per group; Two-way ANOVA with Bonferroni post-hoc test).

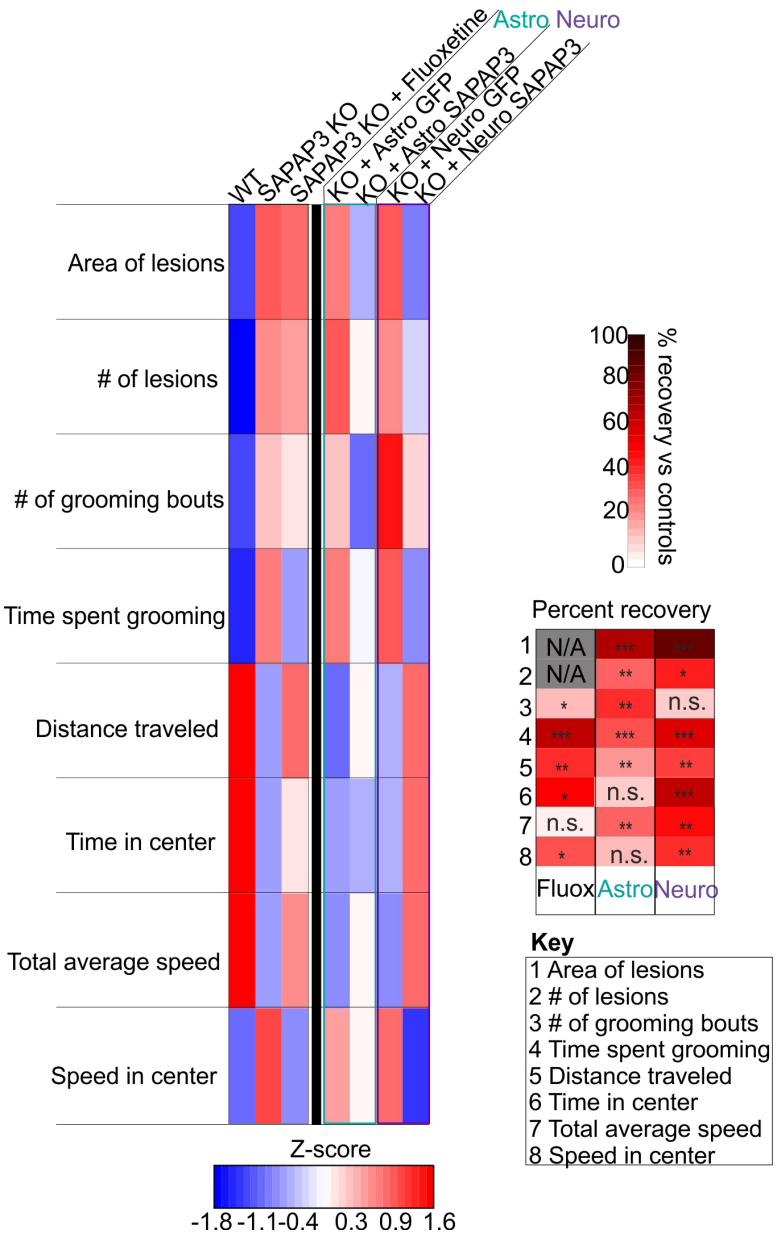

Figure 6.8: Z-score summary of intervention effects on OCD-like behaviors. Heat map shows the z-score for all measured behaviors across intervention with fluoxetine and cell-specific genetic rescue. Subpanel heat maps shows comparison of fluoxetine and cell-specific genetic rescues as a function of percent recovery.

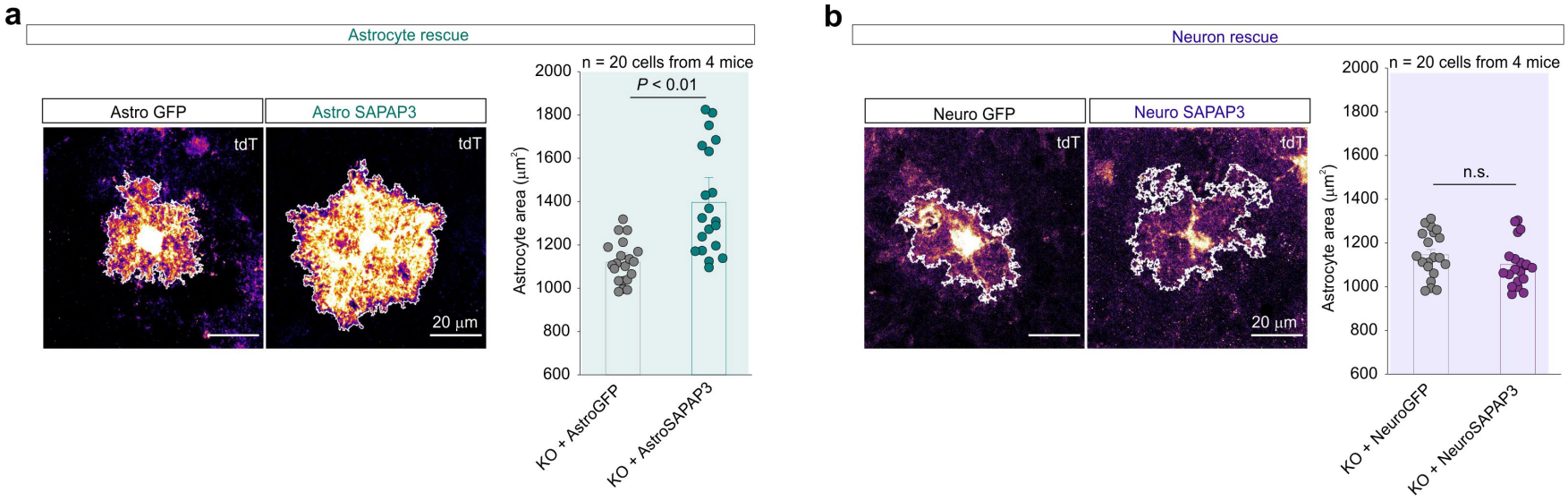

Figure 6.9: Effect of cell-specific genetic rescue on astrocyte morphology. a. Representative images of tdTomato+ striatal astrocytes in SAPAP3 KO mice injected with Astro GFP (left) or Astro SAPAP3 (right). Bar graph shows the measurements of astrocyte territory area. (n = 20 cells from 4 mice per group; two-tailed unpaired t-test with Welch correction, P = 0.001). **b.** Representative images of tdTomato+ striatal astrocytes in SAPAP3 KO mice injected with Neuro GFP (left) or Neuro SAPAP3 (right). Bar graph shows the measurements of astrocyte territory area. (n = 20 cells from 4 mice per group; two-tailed unpaired t-test with Welch correction).

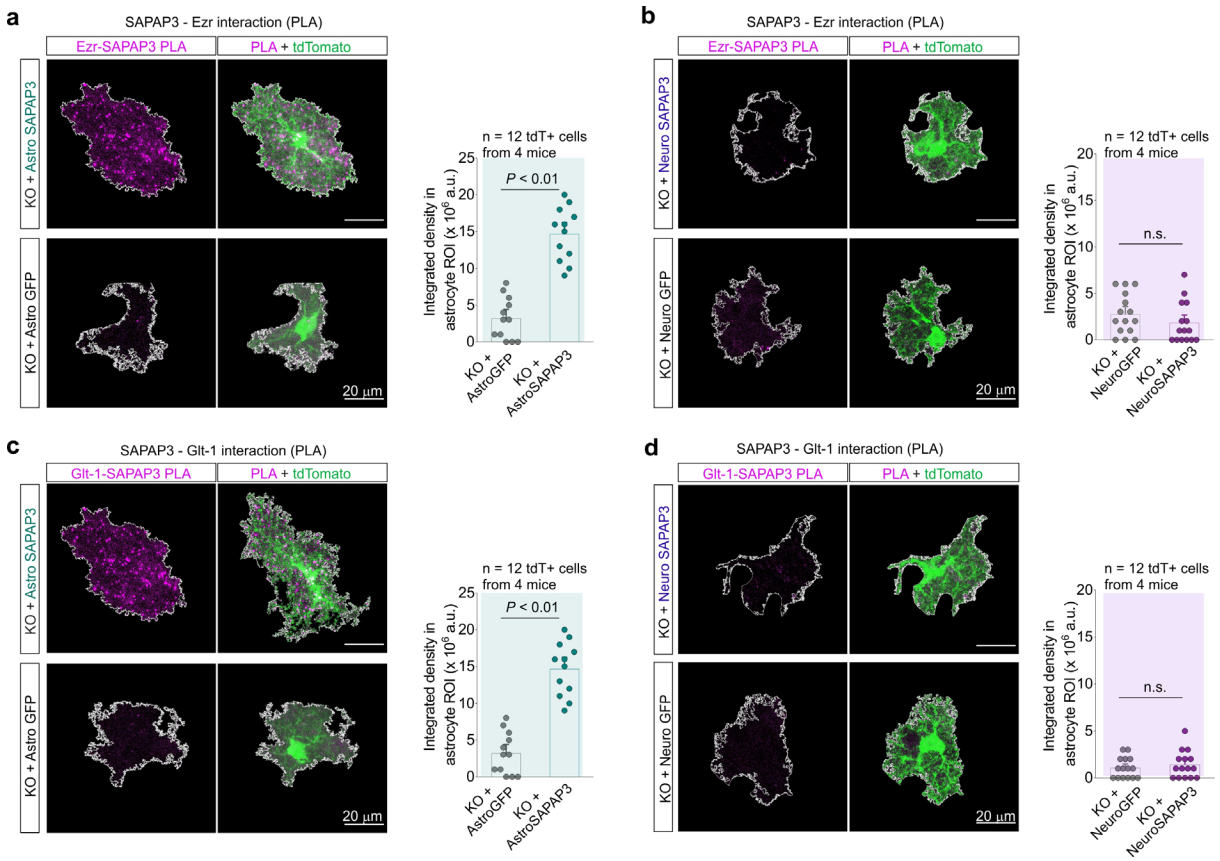

Figure 6.10: Effect of cell-specific genetic rescue on astrocytic SAPAP3 interactions. a. Representative images of PLA puncta for SAPAP3 and Ezrin in SAPAP3 KO mice injected with Astro GFP or Astro SAPAP3. Bar graph denotes the integrated density in arbitrary units (a.u.) of the PLA signal within the astrocyte ROI. (n = 12 tdTomato+ astrocytes from 4 mice per group; two-tailed unpaired t-test with Welch correction, P = 0.002).). **b.** Representative images of PLA puncta for SAPAP3 and Glt-1 in SAPAP3 KO mice injected with Astro GFP or Astro SAPAP3. Bar graph denotes the integrated density in arbitrary units (a.u.) of the PLA signal within the astrocyte ROI. (n = 12 tdTomato+ astrocytes from 4 mice per group; two-tailed unpaired t-test with Welch correction, P < 0.001).). **c.** as in **a**, but for mice treated with Neuro GFP or Neuro SAPAP3. **d.** as in **b**, but for mice treated with Neuro GFP or Neuro SAPAP3.

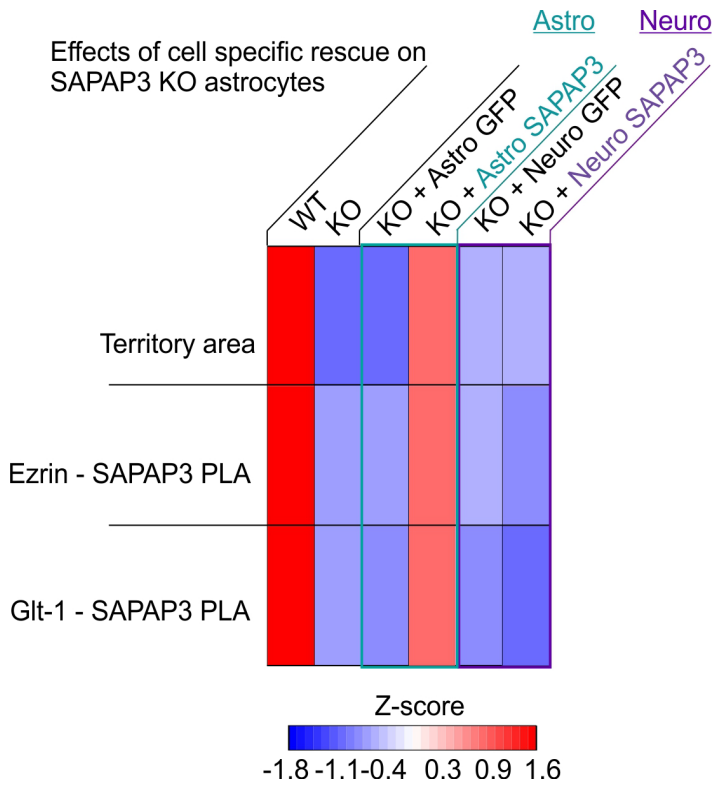

Figure 6.11: Summary of effect of cell-specific genetic rescue on astrocytic SAPAP3 interactions. Heat map shows the Z-scores of the molecular parameters measured in striatal astrocytes in (i) WT and SAPAP3 KO mice, (ii) SAPAP3 KO mice injected with Astro GFP or with Astro SAPAP3, and (iii) SAPAP3 KO mice injected with Neuro GFP or with Neuro SAPAP3.

morphology of the astrocyte via tdTomato signal at ~P180 (6 months of age). Consistent with the behavioral rescue, Astro SAPAP3 significantly restored the astrocyte area size when compared to control Astro GFP (**Figure 6.9**). When we assessed tdTomato+ astrocytes in SAPAP3 KO treated with Neuro SAPAP3, we found no difference in astrocyte area when compared to Neuro GFP as the astrocytes still displayed smaller area sizes (**Figure 6.9**). Additionally, genetic intervention with Astro SAPAP3 restored the both the disrupted astrocytic SAPAP3-Ezrin interaction and the disrupted astrocytic SAPAP3-Glt1 interaction as measured by proximity ligation assay puncta (PLA) within astrocytes (**Figure 6.10**). However, genetic intervention with Neuro SAPAP3 did not restore either SAPAP3-Ezrin or SAPAP3-Glt1 interactions within astrocytes. Taken together, these data highlight the cell-specific mechanisms that drive astrocyte and neuronal contributions to OCD phenotypes in SAPAP3 KO mice. Furthermore, they exemplify that the astrocytic mechanisms driving astrocyte contributions to repetitive behaviors may be driven by cell-autonomous mechanisms that regulate cell morphology and glutamate homeostasis (**Figure 6.11**).

Section 6.5: Effects of cell-specific rescue in neurons and the CSTC circuit

Our results provide evidence that a cell- and striatum-specific intervention can drive profound changes in OCD related behaviors. In order to understand how these changes could arise, we sought to assess the changes that occur in the cortico-striatal-thalamo-cortical (CSTC) circuit which drive these behaviors. As discussed, the CSTC circuit has been heavily implicated in OCD pathophysiology in human studies (Rauch et al., 1997; Saxena and Rauch, 2000; Maia et al., 2008) and circuit dissection of the CSTC circuit in mouse models has revealed that perturbations in specific regions of the circuit can give rise to repetitive, compulsive, and habitual behaviors. Furthermore, cortical inputs onto the striatum, especially from the orbitofrontal cortex, have been

shown to drive repetitive behaviors (Burguiere et al., 2013; Ahmari et al., 2013; Gremel et al., 2013). Therefore, to understand the role of a striatal-specific intervention, it is crucial to distinguish the two parallel pathways that exist within the striatum and drive CSTC function: the direct (D1) and indirect (D2) medium spiny neuron pathways. In brief, the direct (D1) pathway consists of MSNs that express dopamine D1 receptors and project to the internal part of the globus pallidus (GPi) to enhance excitatory feedback to the cortex from the thalamus. Conversely, the indirect pathway (D2) consists of MSNs that express dopamine D2 receptors that reduce feedback to the cortex by projecting to the external part of the globus pallidus (GPe) (Kreitzer et al., 2009).

To assess the changes that occur in the D1 and D2 pathways of the CSTC circuit in a broad and systematic manner, we employed the use of ΔFosB, a well-characterized marker of chronic neuronal activity (**Figure 6.12**). We first assessed ΔFosB levels in WT, SAPAP3 KO, KO + Astro SAPAP3, KO + Astro GFP, KO + Neuro SAPAP3, and KO + Neuro GFP mice in cortical inputs to the striatum including the orbitofrontal cortex (OFC) and the secondary area of the motor cortex (M2). We probed neurons in the OFC and M2 motor cortex with ΔFosB coupled with NeuN neuronal immunostaining. Our results indicated that the number of neurons expressing ΔFosB was significantly higher in SAPAP3 KO mice compared to WT (**Figure 6.13**). This high expression in cortical regions was maintained even in the presence of Astro SAPAP3 and Neuro SAPAP3 genetic rescue conditions suggesting that the behaviorally beneficial effects of cell-specific SAPAP3 rescue are driven by the striatum.

When we assessed local changes of ΔFosB broadly across the striatum, SAPAP3 KO neurons had significantly higher ΔFosB expression than WT (**Figure 6.14**). This was true across the entire striatum as both dorsal and ventral striatum had an equivalent degree of increased ΔFosB+ neurons. Interestingly, in the presence of local SAPAP3 rescue in either astrocytes or neurons,

both dorsal and ventral striatum neurons show significantly less ΔFosB expression when compared to corresponding GFP controls, indicating that local genetic rescue in either astrocytes or neurons produces a direct effect on chronic neuronal activity in the striatum that leads to downstream behavioral consequences. Because of this, we decided to probe whether ΔFosB expression was different in D1 and D2 pathways.

We explored ΔFosB expression in D1 and D2 pathways by using RNA-scope *in situ* hybridization for D1 and D2 receptors coupled with ΔFosB and NeuN immunostaining in the striatum. When we compared ΔFosB+ D1 or D2 MSNs in WT and SAPAP3 KO, we found that both types of MSNs display significantly elevated ΔFosB expression in the KO context (**Figure 6.15**) indicating that both pathways are altered in SAPAP3 KO mice. Upon genetic rescue with Astro SAPAP3, both D1 and D2 MSNs show significantly less expression of ΔFosB when compared to GFP controls. However, rescue with Neuro SAPAP3 only significantly decreases ΔFosB expression in D1 direct pathway neurons with no change in D2 indirect pathway neurons when compared to GFP controls. To further validate these findings, we assessed ΔFosB expression in select nuclei of the CSTC circuit that are downstream of the striatum and are part of the D1 and D2 pathways: the GPe, the GPi, the SNr, and the thalamus. Comparison of WT and SAPAP3 KO mice revealed that both GPi and thalamic neurons express significantly higher levels of ΔFosB in the KO condition (**Figure 6.16**). However, there were no elevated levels of ΔFosB+ neurons in the GPe or SNr in SAPAP3 KO mice. Moreover, Astro SAPAP3 rescue in the striatum caused a significant reduction of ΔFosB+ neurons in the GPe when compared to Astro GFP, but did not cause any changes in the GPi, SNr, nor in the thalamus. However, Neuro SAPAP3 rescue in the striatum caused a significant reduction of ΔFosB+ neurons

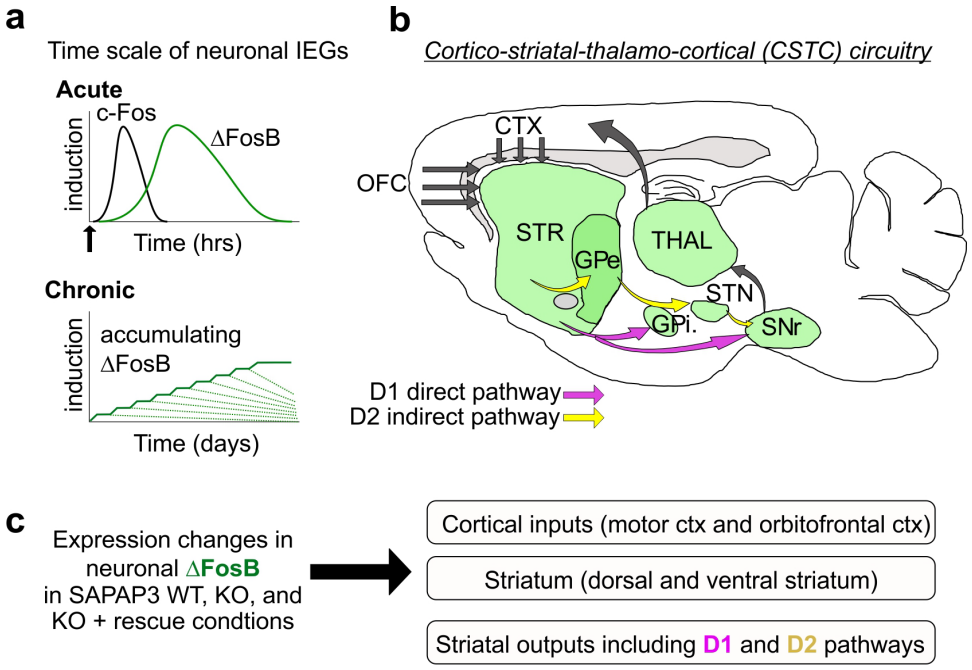

Figure 6.12: Schematic of ΔFosB and the cortico-striatal-thalamo-cortical loop. a. Schematic depicting the time scale of c-Fos and ΔFosB aling with their temporal interaction. **b.** Cartoon showing the cortico-striatal-thalamo-cortical brain pathway and corresponding major D1 and D2 outputs. CTX, Cortex; OFC, orbitofrontal cortex; STR, striatum; GPe, globus pallidus externa; GPi, globus pallidus interna; SNr, substantia nigra; STN, subthalamic nucleus; THAL, thalamus. **c.** Schematic depicting experimental plan for assessment of chronic neuronal changes in SAPAP3 KO mice and different interventions.

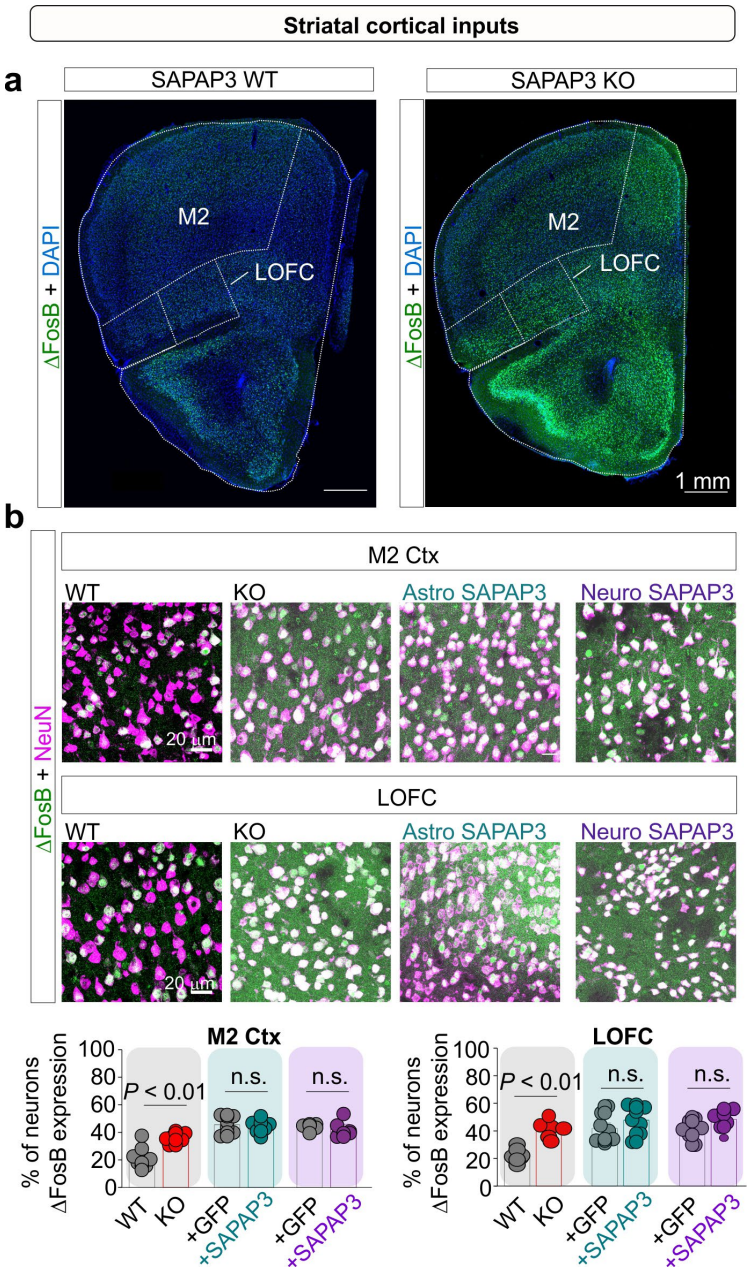

Figure 6.13: ΔFosB changes in cortical inputs to the striatum. a. Representative coronal section images showing ΔFosB expression in wild-type and SAPAP3 KO mice. **b.** Zoom-in images show ΔFosB expression in cortical areas of wild-type, SAPAP3 KO, and SAPAP3 KO with genetic intervention mice. Graphs show the percent of neurons with ΔFosB expression of wild-type, SAPAP3 KO, and SAPAP3 KO with genetic intervention mice (n= 4-5 mice per group; one-way ANOVA with Bonferroni post hoc test).

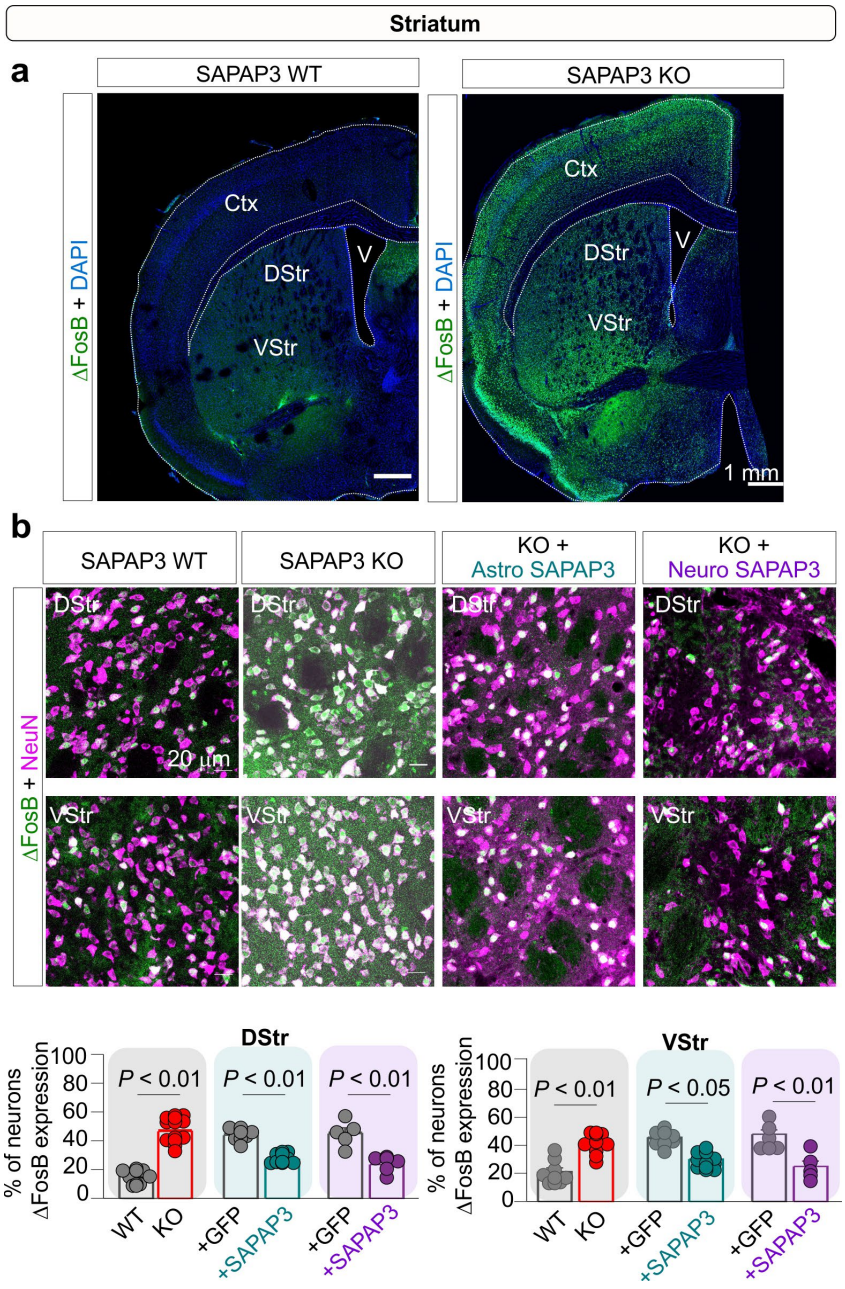

Figure 6.14: ΔFosB changes in the striatum. a. Representative coronal section images showing ΔFosB expression in wild-type and SAPAP3 KO mice. **b.** Zoom-in images show ΔFosB expression in dorsal and ventral striatum of wild-type, SAPAP3 KO, and SAPAP3 KO with genetic intervention mice. Graphs show the percent of neurons with ΔFosB expression of wild-type, SAPAP3 KO, and SAPAP3 KO with genetic intervention mice (n= 4-5 mice per group; one-way ANOVA with Bonferroni post hoc test).

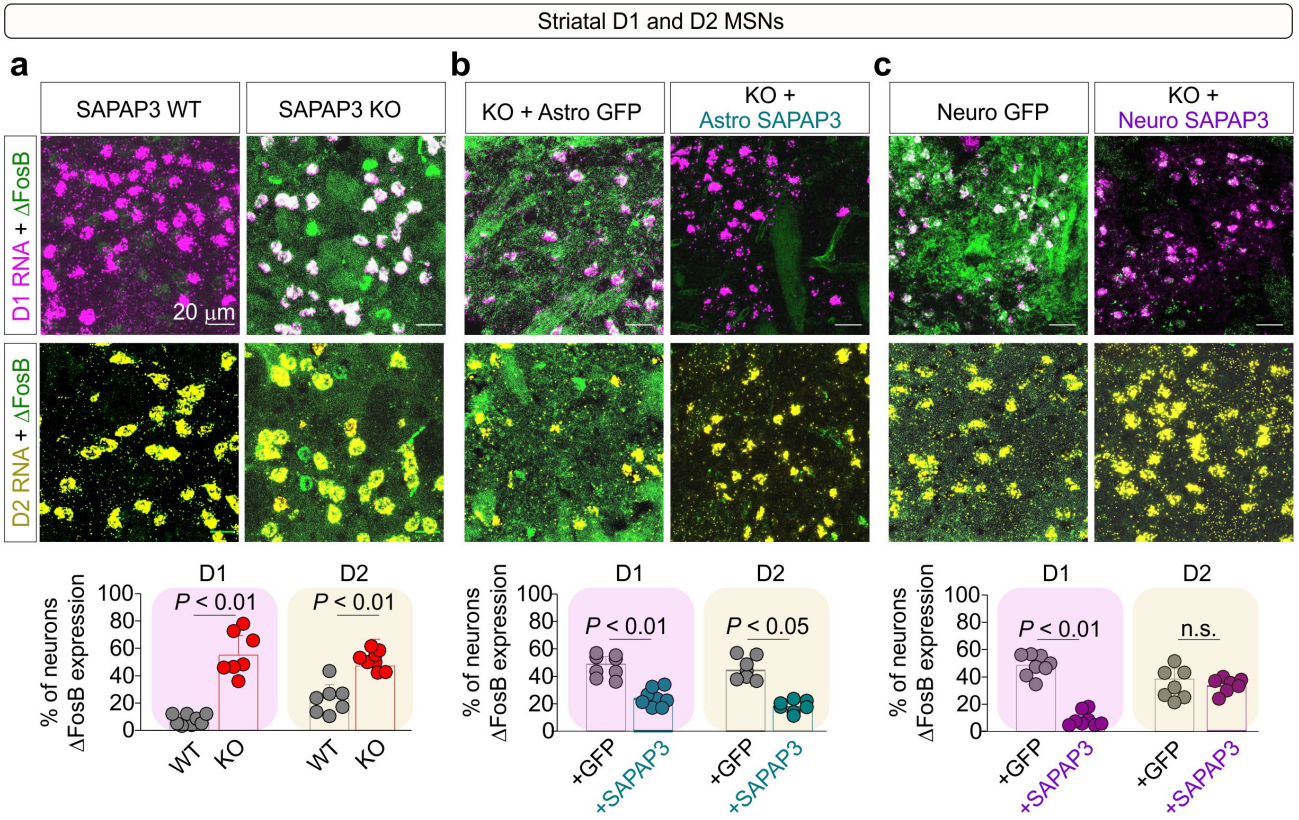

Figure 6.15: ΔFosB changes in D1 and D2 MSNs. a. Representative images of ΔFosB expression in D1 and D2 MSNs of wild-type and SAPAP3 KO mice. Graphs show the percent of neurons with ΔFosB expression of wild-type and SAPAP3 KO mice. (n = 7 mice; One way ANOVA with Bonferroni post-hoc). **b.** Representative images of ΔFosB expression in D1 and D2 MSNs of SAPAP3 KO mice injected with either Astro GFP or Astro SAPAP3. Graphs show the percent of neurons with ΔFosB expression. (n = 6-7 mice; One way ANOVA with Bonferroni post-hoc). **c.** Representative images of ΔFosB expression in D1 and D2 MSNs of SAPAP3 KO mice injected with either Neuro GFP or Neuro SAPAP3. Graphs show the percent of neurons with ΔFosB expression. (n = 6-7 mice; One way ANOVA with Bonferroni post-hoc).

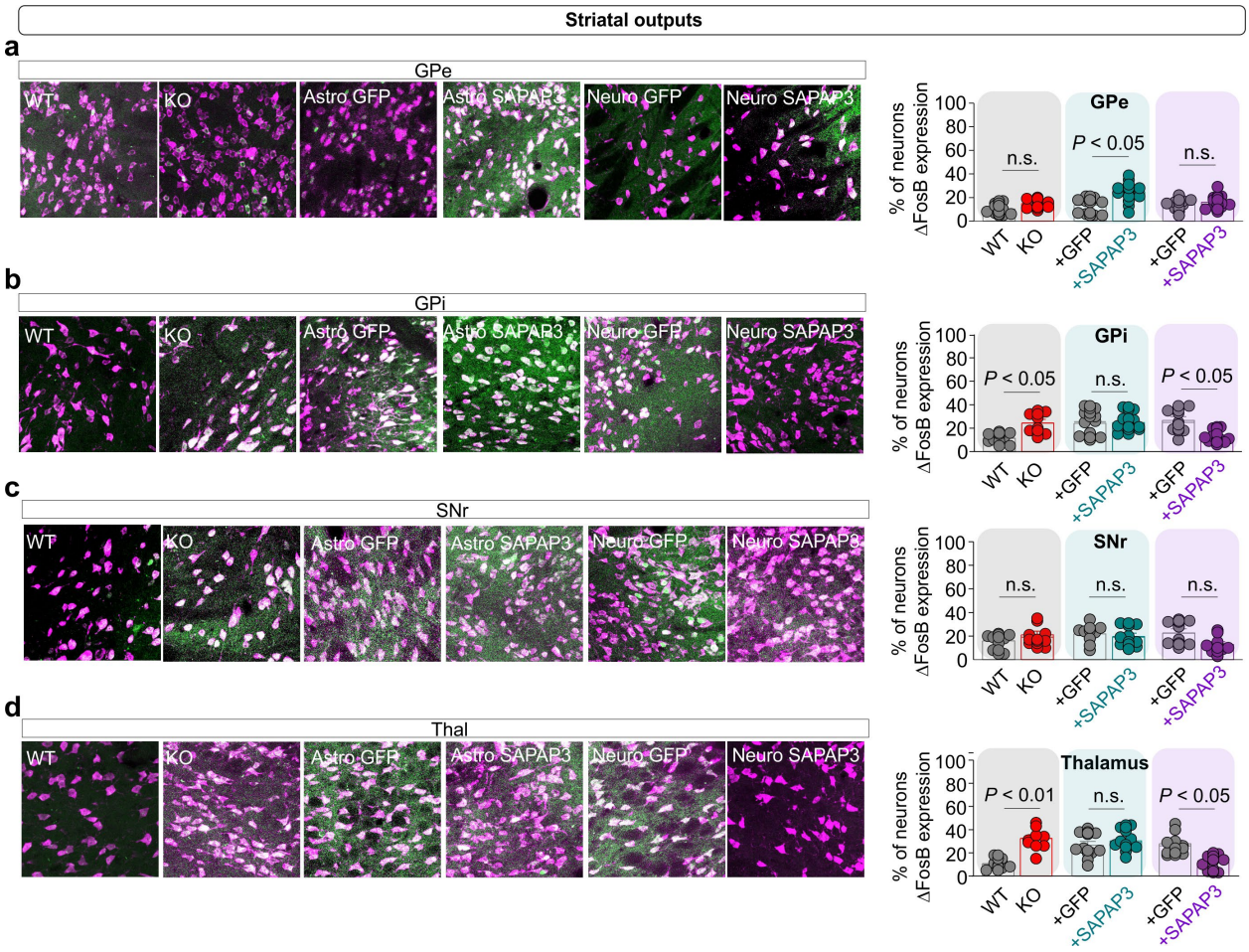

Figure 6.16: ΔFosB changes in striatal output nuclei. a. Representative images of ΔFosB expression in the globus pallidus external segment of wild-type, SAPAP3 KO, and SAPAP3 KO with genetic intervention mice. Graphs show the percent of neurons with ΔFosB expression (n = 4 mice; One way ANOVA with Bonferroni post-hoc). **b.** Representative images of ΔFosB expression in the globus pallidus internal segment of wild-type, SAPAP3 KO, and SAPAP3 KO with genetic intervention mice. Graphs show the percent of neurons with ΔFosB expression (n = 4 mice; One way ANOVA with Bonferroni post-hoc). **c.** Representative images of ΔFosB expression in the substantia nigra of wild-type, SAPAP3 KO, and SAPAP3 KO with genetic intervention mice. Graphs show the percent of neurons with ΔFosB expression (n = 4 mice; One way ANOVA with Bonferroni post-hoc). **d.** Representative images of ΔFosB expression in the thalamus of wild-type, SAPAP3 KO, and SAPAP3 KO with genetic intervention mice. Graphs show the percent of neurons with ΔFosB expression (n = 4 mice; One way ANOVA with Bonferroni post-hoc).

in the GPi and most notably, the thalamus compared to Neuro GFP controls (**Figure 6.16**). This was consistent with the result that the D1 MSN hyperactivity was significantly reversed by Neuro SAPAP3. In all, these data reveal that the differences in behavior between astrocyte and neuronal genetic interventions in the SAPAP3 KO mouse model may arise by specialized signaling mechanisms that modulate D1 and D2 pathways distinctly. Therefore, we provide a framework by which we may explore how astrocytes modulate D1 and D2 pathways and how MSNs respond to regulatory mechanisms that are both cell-autonomous and non-cell autonomous.

Section 6.6: Relevance to human OCD data

The experiments presented here provide an exciting assessment of cell-specific mechanisms that drive repetitive and anxiety behaviors in a mouse model of OCD. In order to evaluate the relevance of our results to human OCD, we used mass spectrometry to conduct bulk proteomics on striatum of WT and SAPAP3 KO mice so that we could assess how protein changes compare to gene expression changes in post-mortem human OCD tissue and to astrocyte and neuronal gene expression. Our bulk proteomics experiment revealed 66 differentially expressed proteins (DEPs), 46 of which were upregulated and 20 that were downregulated in SAPAP3 KO mice when compared to WT (**Figure 6.17**). When we assessed the gene expression level of these 66 DEPs in astrocytes and neurons, we found that all 66 proteins were expressed at the mRNA level (FPKM > 5) in both astrocytes and neurons from our RiboTag experiments (**Figure 6.18**). Furthermore, about 35% of our DEPs were found to be in either astrocytes or neurons at the protein level. We next compared our DEPs to their gene expression level in human OCD post-mortem tissue we found that 44 of the 66 DEPs were dysregulated in human OCD with 18 showing similar directional changes including *Grin2a* and *Ankrd6*.

When we identified the top 30 differentially expressed caudate (human striatum) genes between healthy human non-diseased individuals and OCD individuals (Piantadosi et al., 2021) we found that 29 of the 30 genes were expressed in both astrocytes and neurons (FPKM > 5) and 2 were found within astrocyte and neuronal proteomes (**Figure 6.19**). Finally, we explored astrocytic and neuronal expression for 61 genes associated with human OCD and Tourette's syndrome (Yu et al., 2010), a disease that shares common features with OCD and is in the OCD spectrum. These analyses revealed that these 61 genes were all expressed to some degree in astrocytes and neurons (FPKM > 1) and at least 15% and 40% of them were found within our astrocyte and neuron proteomes, respectively. Additionally, a few were within our cell-specific SAPAP3 putative interactor lists (**Figure 6.19**).

Because a feature of human OCD is the development of symptoms during discrete developmental time points, especially at adolescence, we assessed the expression of striatal SAPAP3 across different mouse developmental time-points postnatally. Surprisingly, we found that postnatal expression of SAPAP3 mRNA in astrocytes and neurons differed in mice. These results revealed that at early developmental time points (P1-P7) SAPAP3 mRNA is exclusively expressed by NeuN+ neurons and not by Sox9+ astrocytes. Sox9+ astrocytes do not begin expressing SAPAP3 until P14. From P14 ,SAPAP3 expression within Sox9+ astrocyte increases and then plateaus at ~P56, while SAPAP3 expression in NeuN+ neurons stays high and constant throughout the entirety of postnatal development (**Figure 6.20**). These results may give a clue of how SAPAP3 expression may be relevant to the emergence of OCD phenotypes during adolescence.

The human analyses conducted here support our results in mice that molecular changes that occur or are associated with human OCD can impact signaling mechanisms in both astrocytes and

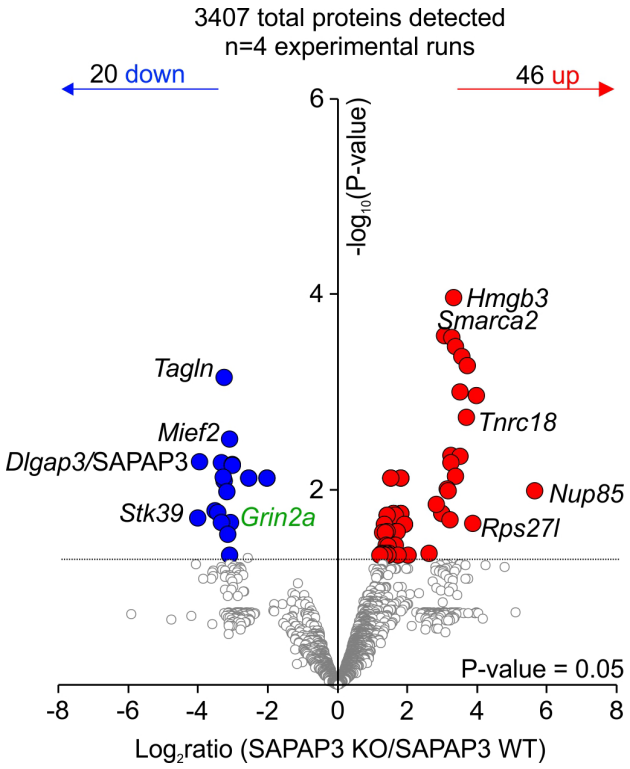

Figure 6.17: Whole tissue striatal proteomics of WT and SAPAP3 KO mice. Volcano plot shows the label-free based quantification comparison of 3407 proteins detected in whole tissue striatum proteomics. Red circles show upregulated proteins (Log2FC > 1 and FDR < 0.05) while blue circles show downregulated proteins (Log2FC < 1, FDR < 0.05) in the SAPAP3 KO striatum when compared to wild-type. The top 5 most down regulated and up regulated proteins are shown. Green label shows protein that appeared in the neuronal SAPAP3 interactome

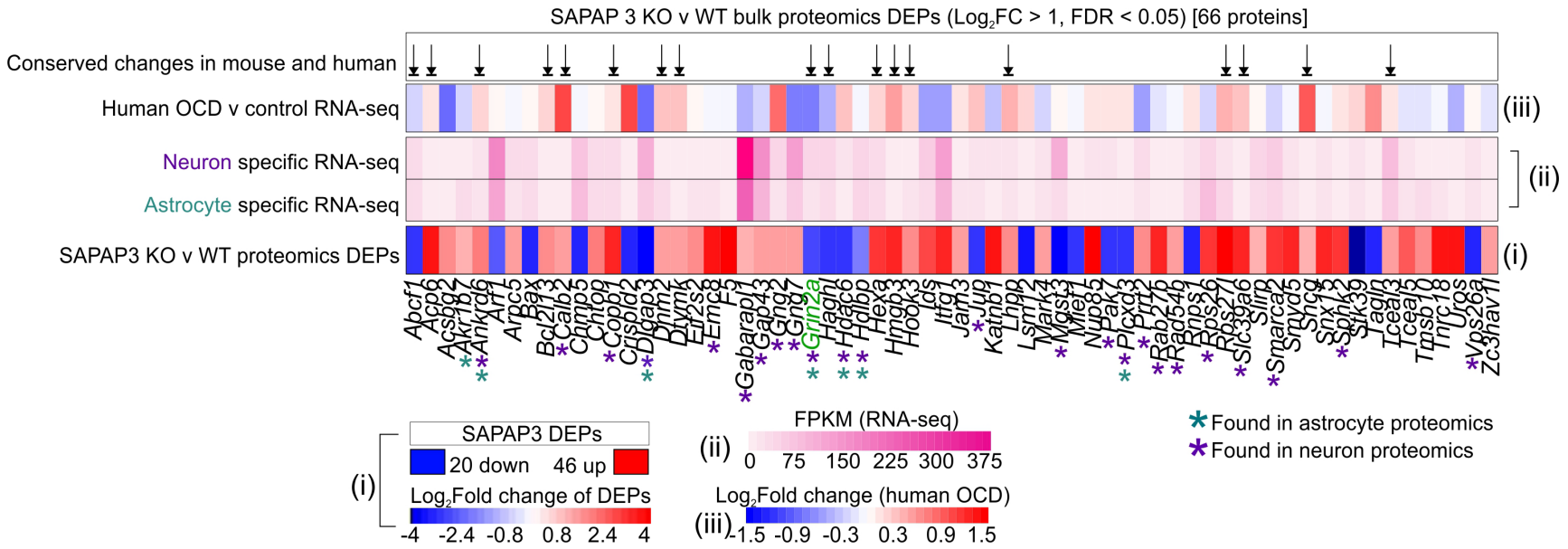

Figure 6.18: Comparison of mouse striatal proteomics data to human OCD RNA-seq. List of the 66 differentially expressed proteins in the striatum of SAPAP3 KO mice when compared to wild typ. Heat map (*i*) shows the Log_2fold change of the 66 proteins *versus* wild type control. Heat map (*ii*) shows the mRNA abundance (FPKM) of the 66 differentially expressed proteins in our neuron or astrocyte specific mouse RiboTag RNA-seq datasets. Heat map (*iii*) shows the Log_2fold change at the mRNA level of the 66 proteins in human caudate of OCD subjects compared to controls. Arrows show genes that had conserved changes in the mouse SAPAP3 KO model and human OCD. Teal asterisks denote whether the protein was found in the astrocyte specific proteomics datasets, while the purple asterisks denote whether the protein was found in the neuron specific proteomics datasets.

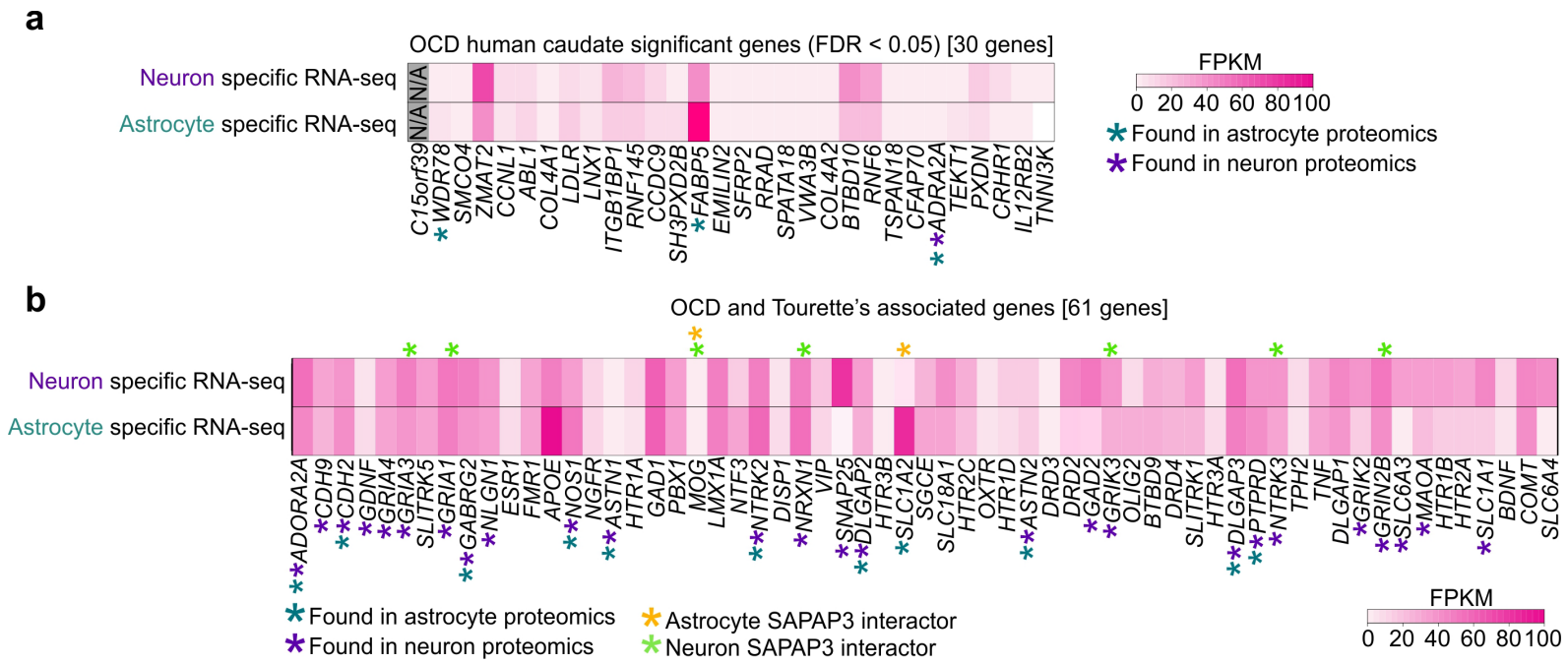

Figure 6.19: Assessment of human OCD-related genes to cell-specific RNA-seq and proteomics. List of 61 genes associated with OCD and Tourette's syndrome from Phenopedia. Heat map depicts the genes' respective mRNA abundances (FPKM) in our mouse neuron or astrocyte specific RNA-seq datasets. Teal asterisks show whether the protein was also found in the astrocyte specific proteomics datasets while the purple asterisks show whether the protein was found in the neuron specific datasets. Orange asterisks denote whether the protein was an astrocytic SAPAP3 interactor, while green asterisks denote whether the protein was a neuronal SAPAP3 interactor.

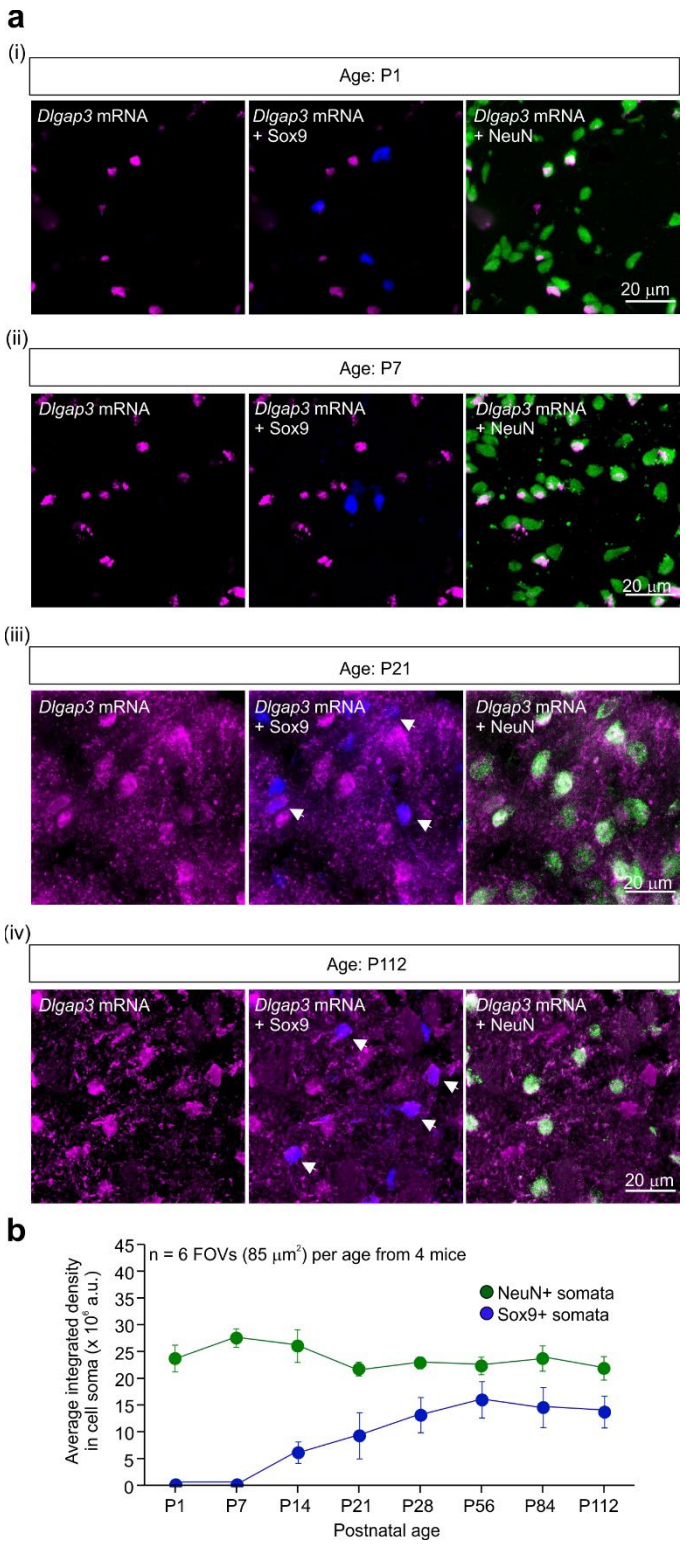
a
(i)
Age: P1
Dlgap3 mRNA
Dlgap3 mRNA + Sox9
Dlgap3 mRNA + NeuN
20 μm
(ii)
Age: P7
Dlgap3 mRNA
Dlgap3 mRNA + Sox9
Dlgap3 mRNA + NeuN
20 μm
(iii)
Age: P21
Dlgap3 mRNA
Dlgap3 mRNA + Sox9
Dlgap3 mRNA + NeuN
20 μm
(iv)
Age: P112
Dlgap3 mRNA
Dlgap3 mRNA + Sox9
Dlgap3 mRNA + NeuN
20 μm
b
n = 6 FOVs (85 μm²) per age from 4 mice
NeuN+ somata
Sox9+ somata
Average integrated density in cell soma (X 10² a.u.)
45
40
35
30
25
20
15
10
5
0
P1
P7
P14
P21
P28
P56
P84
P112
Postnatal age

Figure 6.20: SAPAP3 mRNA expression across postnatal development. a. (i-iv) Representative images of P1-P112 mouse striata with SAPAP3 mRNA in magenta, Sox9 astrocyte marker protein in blue, and NeuN neuronal marker protein in green. White arrows show SAPAP3 mRNA expression in Sox9+ astrocytes. **b.** Line graph shows the average integrated intensity of SAPAP3 mRNA in either NeuN+ somata (green) or Sox9+ somata (blue) across 8 postnatal ages. Data are presented as the mean and the SEM; n = 6 FOVs per age from 4 mice per age.

neurons. Furthermore, we provide a handful of exciting new directions by which we can explore relevant mechanisms of OCD in a cell-specific manner.

Section 6.7: Discussion

The contributions of astrocytes and neurons to brain disorders, including neuropsychiatric disease, have long been speculated (Kelley et al., 2018; Yu et al., 2018; Cui et al., 2018). Here, we explored novel astrocyte and neuronal mechanisms relevant to OCD that were directed by our broad proteomic explorations of astrocyte and neuronal function in the striatum. OCD is a poorly treated and complex disease characterized by obsessive thoughts that become intrusive and lead to compulsions exhibited as repetitive behaviors with accompanied anxiety. While OCD involves cortico-striatal circuit dysfunction, the precise cellular and molecular mechanisms that drive the disorder remain unclear. Previous studies have demonstrated that diverse cell types contribute to repetitive behaviors (Chen et al., 2010; Yu et al., 2018), indicating multicellular contributions to circuit dysfunction that ultimately give rise to OCD-like phenotypes. Our results demonstrate that both astrocyte and neuron SAPAP3 mechanisms contribute to OCD phenotypes in mice and that these mechanisms are shared and distinct at both the molecular and behavioral level. Thus, our experiments indicate that therapeutic strategies that target multicellular mechanisms could be useful to explore in a range of brain disorders, including OCD.

CHAPTER 7: CONSEQUENCES OF ASTROCYTIC CHEMOGENETIC MODULATION IN A MOUSE MODEL OF OBSESSIVE-COMPULSIVE DISORDER

Section 7.1: Introduction

Using broad exploratory proteomic methods in both astrocytes and neurons of the striatum, we discovered that both cell types express SAPAP3, a gene related to obsessive-compulsive disorder (**Chapters 3, 4 and 5**). Our explorations also demonstrated that astrocytes express a rich variety of G-protein coupled receptors (GPCRs), a large family of membrane proteins that mediate most cellular responses to extracellular cues by integrating and amplifying intracellular signaling cascades (Rosenbaum et al., 2009). Furthermore, when we explored astrocyte specific mechanisms of SAPAP3 we found that four of the SAPAP3 interactors were related to G-protein signaling, including Gnai2 (Gene: *Gnai2*) a G-protein that directly integrates signals from GPCRs and initiates the intracellular signal cascade known as the G(i) pathway.

Because GPCR pathways mediate complex signaling events that give rise to cellular responses to a variety of extracellular stimuli, they have the potential to become important therapeutic targets for a variety of diseases. Furthermore, the expression of different subtypes of GPCRs in different cell types allow for exploration of cell-specific biology using pharmacological or engineered methods. DREADDs (designer receptors exclusively activated by designer drugs) are engineered GPCRs that have little to no responses to their endogenous ligands and have been designed to respond to synthetic and biologically inert chemical ligands (Roth 2016). As such, DREADDs have been widely used to explore neuron and astrocyte biology in brain preparations and *in vivo* by genetic expression of the DREADDs in distinct brain circuits with subsequent systemic injection of the synthetic ligand. Thus, DREADDs allow activation of distinct G-protein pathways

including the G(i) pathway via hM4D DREADD, the G(q) pathway via hM3D DREADD, and the G(s) pathway via rM3D DREADD in a time-dependent and cell-specific manner.

In relation to astrocyte biology, several studies have provided links between GPCR pathways within astrocytes to rodent behavioral alterations relevant to psychiatric disorders. For example, activation of the G(i)-GPCR signaling pathway in striatal astrocytes with hM4Di DREADD resulted in behavioral phenotypes related to hyperactivity with disrupted attention caused by enhanced excitatory transmission onto medium spiny neurons (MSNs) (Nagai et al., 2019). Additionally, activation of the G(i)-GPCR signaling pathway in striatal astrocytes with hM4Di DREADD improved detrimental behavioral phenotypes in the R6/2 mouse model of Huntington's disease (Yu et al., 2020). In the context of the R6/2 mouse model of Huntington's disease, astrocytes displayed reduced territory sizes which were also improved by G(i)-GPCR activation via hM4Di (Yu et al., 2020). Furthermore, genetic attenuation of the G(q)-GPCR signaling pathway within astrocytes *in vivo* has direct consequential effects on mouse behaviors including impaired spatial memory and altered neuronal signaling (Nagai et al., 2021). Thus, astrocyte specific G-protein pathways reveal causal astrocytic mechanisms that modulate circuits and therefore, behavior.

To further assess how astrocyte specific signaling cascades are involved in SAPAP3 mediated mechanisms that give rise to OCD-like phenotypes, we studied the consequences of G(i) and G(q) activation in striatal astrocytes using DREADDs in SAPAP3 KO mice. Our goal was to evaluate whether activation of these G-protein pathways provided any effect on the compulsion and anxiety behaviors in SAPAP3 KO mice and whether we could map these changes to neuronal activity in the striatum. In this way, specific molecular mechanisms could be teased apart to design astrocyte GPCR-based strategies to modulate responses related to psychiatric diseases such as OCD.

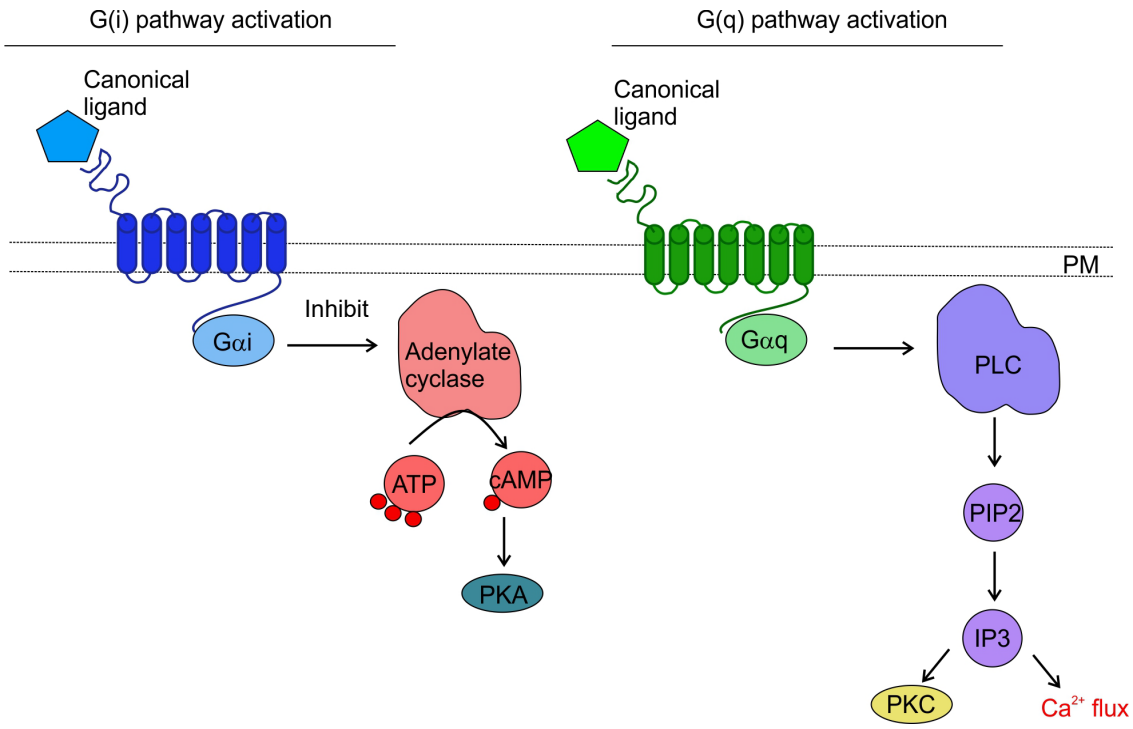

Figure 7.1: Canonical molecular pathways of G(i) and G(q) G-signaling. Cartoon shows the canonical molecular signaling pathways of G(i) and G(q) G-protein signaling pathways upon ligand activation of a GPCR.

Section 7.2: Consequences of striatal astrocyte Gq pathway activation on OCD-like behaviors

Because both G(i) and G(q) GPCR pathway modulation in astrocytes has been shown to be consequential to striatum associated behaviors, we first assessed the consequences of striatal astrocyte G(q) activation on both compulsion and anxiety behaviors in SAPAP3 KO mice. To further elaborate and distinguish the pathways, when a ligand binds to a G(q) associated GPCR, the intracellular Gq G-protein dissociates from the GPCR and interacts with enzymes of the phospholipase C family (PLC). These enzymes in turn regulate Ca^{2+} flux and activation of a variety of second messengers which cause a variety of intracellular signaling events to occur (**Figure 7.1**). To study striatal astrocyte G(q) pathway activation in the SAPAP3 KO context, we employed the use of astrocyte specific hM3Dq DREADD AAV (AAV 2/5 *gfaABC$_1$D* hM3Dq-mCherry), hereafter referred to as Gq. We microinjected either Gq or astrocyte specific tdTomato as a control bilaterally into the striatum of wild-type (WT) and SAPAP3 KO mice at ~P28. After one month of expression, we injected deschloroclozapine (DCZ), a DREADD specific agonist, systemically into all mice via intraperitoneal (i.p.) injection (10μg/kg) three times a week for three months. We had four total experimental groups: 1) WT + tdTomato, 2) WT + Gq, 3) SAPAP3 KO + tdTomato, and 4) SAPAP3 KO + Gq which were all treated with DCZ. At ~P180, or about 6 months of age, we preformed behavioral assessments related to compulsion and anxiety including self-grooming, lesion analysis, and elevated plus maze (**Figure 7.2**).

Ambulation assessments in the open field arena (OFA) showed that astrocytic Gq activation via hM3D DREADD had no effect in the WT mice as both tdT and Gq groups both had the same average total speed and travelled the same distance in the OFA. When we assessed the SAPAP3 KO mice, both tdT and Gq groups showed decreased total speed and total distance traveled when

compared to WT. Comparison between KO + tdT and KO + Gq groups did not show any difference in ambulation as both groups had similar total average speed and total distance traveled measurements. Thus, astrocytic Gq activation did not affect or improve measures of general ambulation in the SAPAP3 KO mice (**Figure 7.3**).

When we assessed measurements of anxiety with the OFA and the elevated plus maze (EPM), we again, found no effect of Gq activation between the WT groups. We also did not find an effect of Gq activation when we compared the KO + tdT and the KO + Gq groups as both groups showed decreased time spent in the open arms of the EPM, increased speed in the center of the OFA, and decreased time spent in the center of the OFA when compared to WT. These metrics did not improve with astrocytic Gq activation (**Figure 7.4**).

Finally, when we explored excessive self-grooming by using the area of lesions, the total number of lesions, the number of grooming bouts, and the time spent grooming, we found no significant effect of astrocytic Gq activation in either WT or SAPAP3 KO in any of these metrics. Comparison of WT + tdT and WT + Gq showed no changes in any grooming parameters and this result was the same for the KO + tdT and KO + Gq group comparison. Thus, activation of the Gq GPCR pathway in astrocytes does not affect self-grooming in the context of SAPAP3 KO (**Figure 7.5**).

Our results indicate that astrocytic activation of Gq in the striatum does not change any compulsion or anxiety phenotypes in the context of OCD. Activation of Gq did not change any baseline behaviors in WT mice nor did they change any OCD-related behaviors in SAPAP3 KO mice. Therefore, activation of the astrocytic Gq pathway is not consequential to OCD-related behaviors in the SAPAP3 KO mouse model.

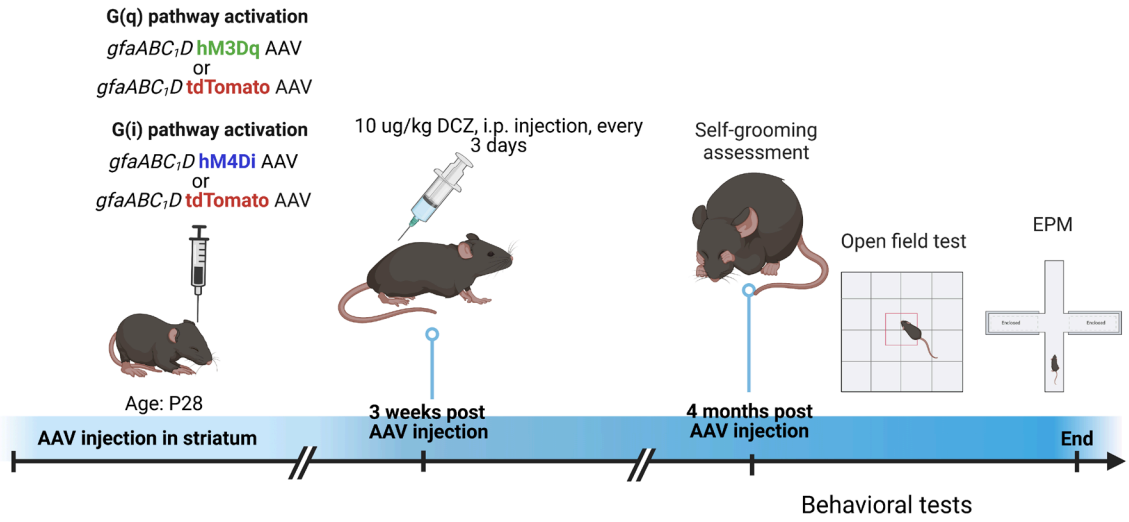

Figure 7.2: Experimental timeline for G(q) and G(i) pathway activation. Cartoon shows the experimental plan and timeline for *in vivo* activation of the astrocytic (Gq) pathway or astrocytic G(i) pathway. Wild-type or SAPAP3 KO mice are microinjected with G(q) or G(i) DREADDs and then chronically treated with the ligand DCZ. Image made with Biorender (Biorender.com).

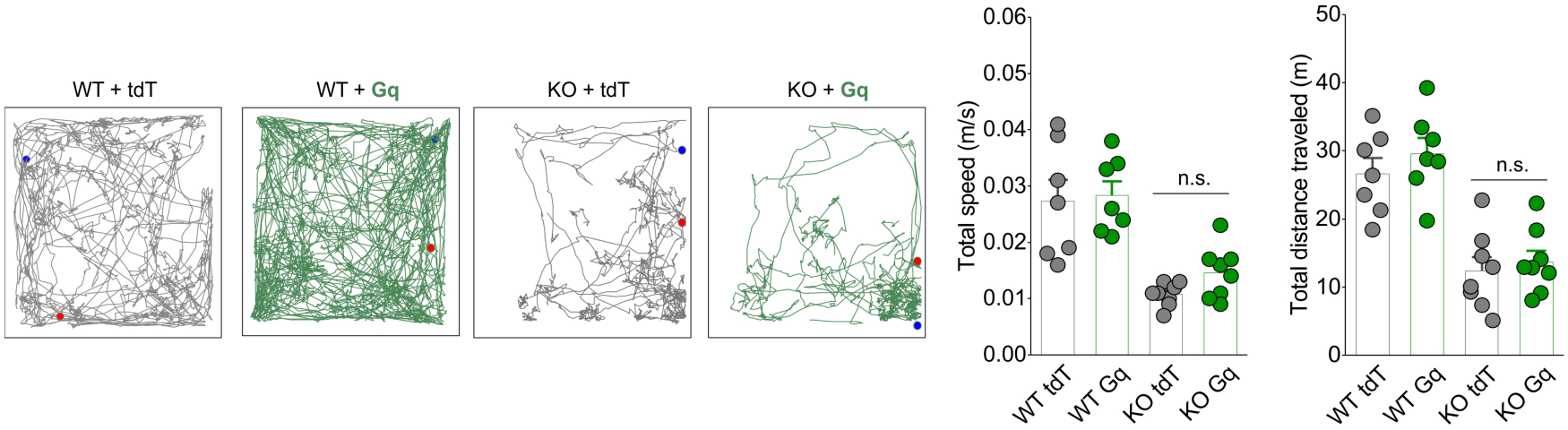

Figure 7.3: No effect of astrocytic G(q) activation on locomotion in SAPAP3 KO mice. Representative traces of 20 minute open field recordings for WT and SAPAP3 KO mice treated with DCZ and microinjected with tdTomato or hM3Dq Gq GPCR. Blue dot depicts the location of the mouse at the start of the recording, while the red dot depicts the location of the mouse at the end of the recording. Bar graphs show locomotor activity by two different metrics: total distance traveled and total average speed. (n = 7-8 mice per group; Kruskal-Wallis test).

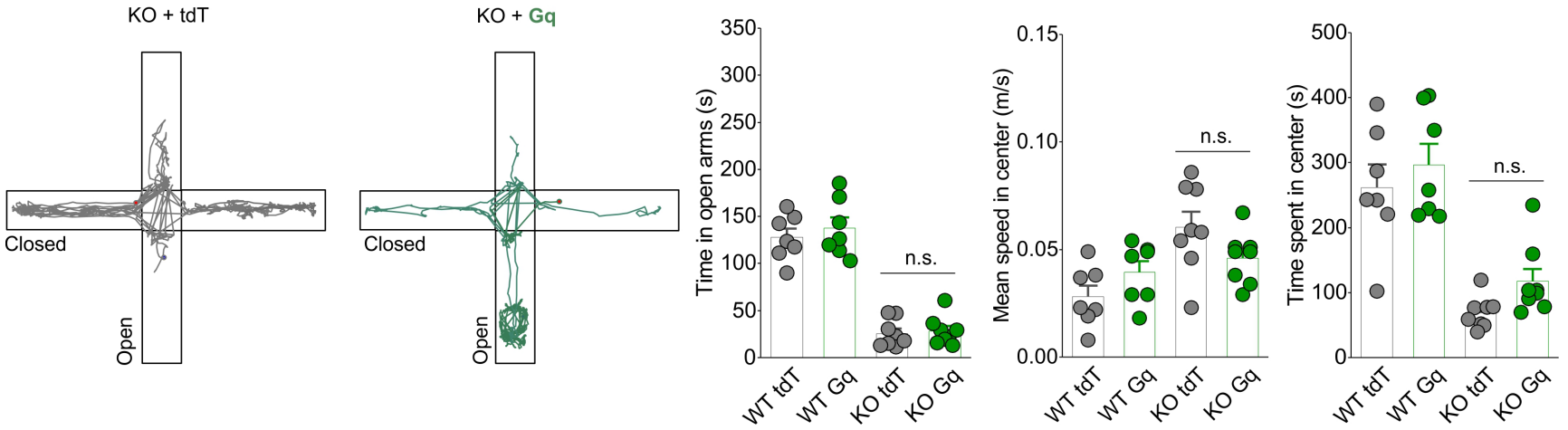

Figure 7.4: No effect of astrocytic G(q) activation on anxiety in SAPAP3 KO mice. Representative traces of 10 minute elevated plus maze recordings for WT and SAPAP3 KO mice treated with DCZ and microinjected with tdTomato or hM3Dq Gq GPCR. Blue dot depicts the location of the mouse at the start of the recording, while the red dot depicts the location of the mouse at the end of the recording. Bar graphs show anxiety parameters measured by three different metrics: Time spent in open arms of the EPM, mean speed in the center of the OFA, and time spent in the center of the OFA. (n = 7-8 mice per group; Kruskal-Wallis test).

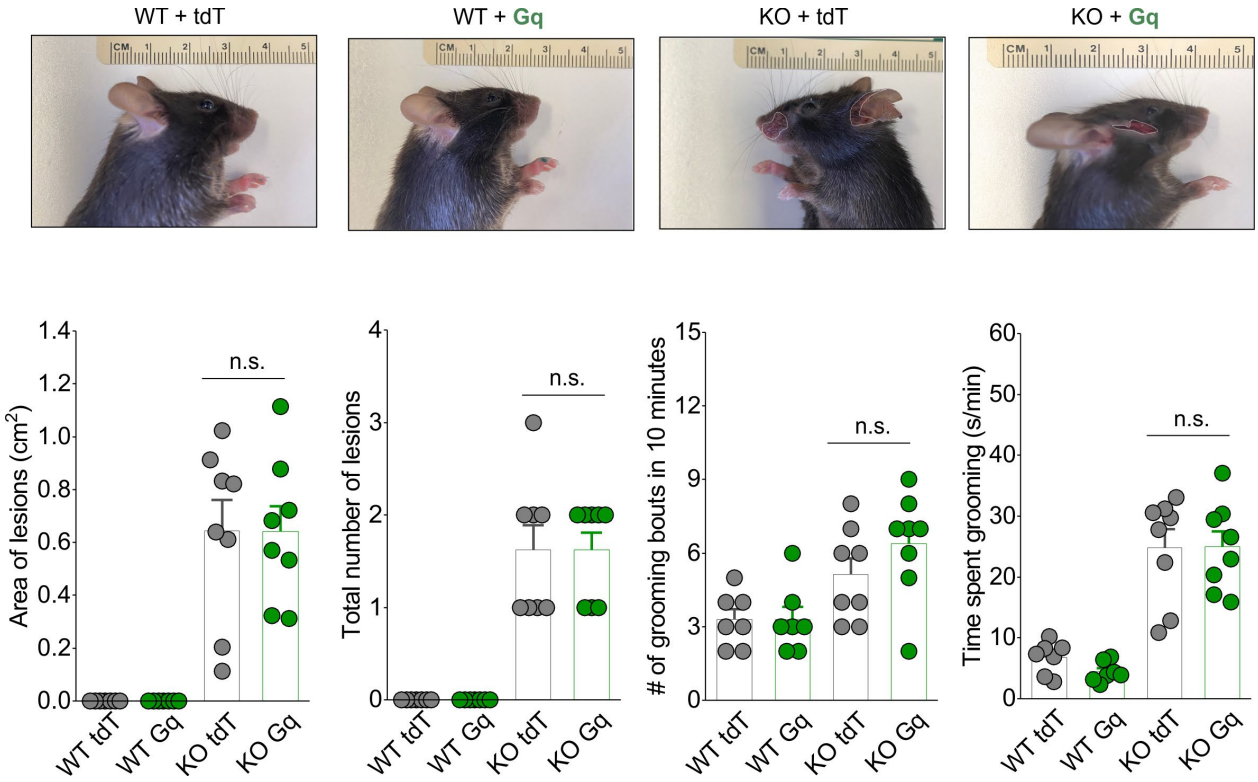

Figure 7.5: No effect of astrocytic G(q) activation on self-grooming in SAPAP3 KO mice. Representative photographs of WT and SAPAP3 KO mice treated with DCZ and microinjected with tdTomato or hM3Dq Gq GPCR. White outlines show the extent of the lesions. Bar graphs show self-grooming parameters measured by four different metrics: area of lesions per mouse, total number of lesions per mouse, number of grooming bouts in 10 minutes, and time spent grooming. (n = 7-8 mice per group; Kruskal-Wallis test).

Section 7.3: Consequences of striatal astrocyte Gi pathway activation on OCD-like behaviors

Our previous explorations revealed that proteins related to the G(i) GPCR pathway, including Gnai2, were constituents of the SAPAP3 interactome within striatal astrocytes. Thus, we assessed the consequences of astrocytic G(i) pathway activation in the context of SAPAP3 KO. In short, when a ligand binds to a G(i) associated GPCR, the intracellular Gi G-protein dissociates from the GPCR and inhibits adenylate cyclase, a conserved enzyme that mediates the conversion of adenosine triphosphate (ATP) to cyclic-AMP (cAMP) which then in turn acts as a second messenger to mediate a myriad of intracellular signaling events. Thus, the G(i) pathway directly regulates the concentration of cAMP (**Figure 7.1**). To explore the role of astrocyte specific G(i) pathway activation in the striatum and in the context of SAPAP3 KO, we used astrocyte specific hM4Di DREADD AAV (AAV 2/5 *gfaABC$_1$D* hM4Di-mCherry), hereafter referred to as Gi. We microinjected wither Gi or astrocyte specific tdTomato as a control bilaterally into the striatum of WT and SAPAP3 KO mice at ~P28. After one month of expression, we began injection of deschloroclozapine (DCZ), a DREADD specific agonist, systemically into all mice intraperitoneally (i.p.) (10μg/kg) for three times a week for three months. We a total of four experimental groups: 1) WT + tdTomato, 2) WT + Gi, 3) SAPAP3 KO + tdTomato, and 4) SAPAP3 KO + Gi which were all treated with DCZ. At ~P180, or about 6 months of age, we preformed behavioral assessments related to compulsion and anxiety including self-grooming, lesion analysis, and elevated plus maze (**Figure 7.2**).

Measurements of ambulation in the OFA revealed that there were no differences in total speed and total distance traveled between the WT + tdT and the WT + Gi groups. However, comparison of the KO + tdT and KO + Gi groups revealed a significant increase of both total speed and total distance traveled in KO mice with astrocytic Gi activation when compared to KO + tdT control.

Our results indicate that activation of the G(i) pathway selectively within striatal astrocytes improves ambulation and locomotion in SAPAP3 KO mice (**Figure 7.6**).

When we assessed anxiety behavior in the elevated plus maze (EPM) and the OFA, we found that WT + Gi mice trended to explore the open arms of the EPM for more time when compared to WT + tdT mice. When we compared the KO + tdT and KO + Gi groups, we found that SAPAP3 KO mice with astrocytic Gi activation explored the open arms significantly more than the respective control mice. Measurements in the OFA also showed that KO + Gi mice spent significantly more time in the center and were significantly slower when crossing the center of the open field, thus indicating less anxiety. Therefore, astrocytic Gi pathway activation improves anxiety in the context of SAPAP3 (**Figure 7.7**).

Finally, we measured several self-grooming parameters to explore consequences of astrocytic Gi pathway activation in WT and SAPAP3 KO mice. We measured the area of self-inflicted lesions, the total number of lesions, the number of grooming bouts, and the time spent grooming. Surprisingly, comparison of the KO + tdT and KO + Gi group revealed significantly decreased area of lesions, significantly reduced number of lesions, and significantly reduced grooming time in KO + Gi mice when compared to KO + tdT mice. The number of grooming bouts remained unchanged. Astrocytic Gi activation had no effect on baseline self-grooming in WT mice. Therefore, as the astrocyte SAPAP3 interactome predicted, activation of the G(i) pathway in striatal astrocytes of SAPAP3 KO mice improves excessive self-grooming in these mice (**Figure 7.8**).

Our results show that specific activation of the G(i) GPCR pathway in striatal astrocytes is sufficient to ameliorate some of the detrimental phenotypes in the SAPAP3 KO model of OCD including excessive self-grooming and anxiety. Interestingly, these results were not re-capitulated

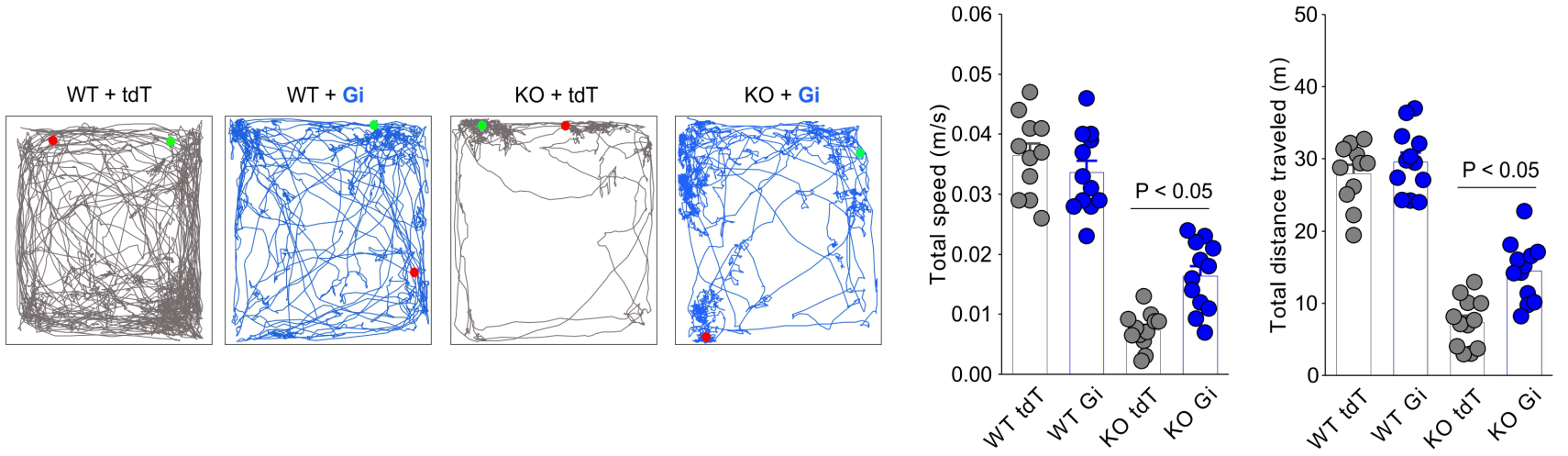

Figure 7.6: Effect of astrocytic G(i) activation on locomotion in SAPAP3 KO mice. Representative traces of 20 minute open field recordings for WT and SAPAP3 KO mice treated with DCZ and microinjected with tdTomato or hM4Di Gi GPCR. Green dot depicts the location of the mouse at the start of the recording, while the red dot depicts the location of the mouse at the end of the recording. Bar graphs show locomotor activity by two different metrics: total distance traveled and total average speed. (n = 11-12 mice per group; Kruskal-Wallis test).

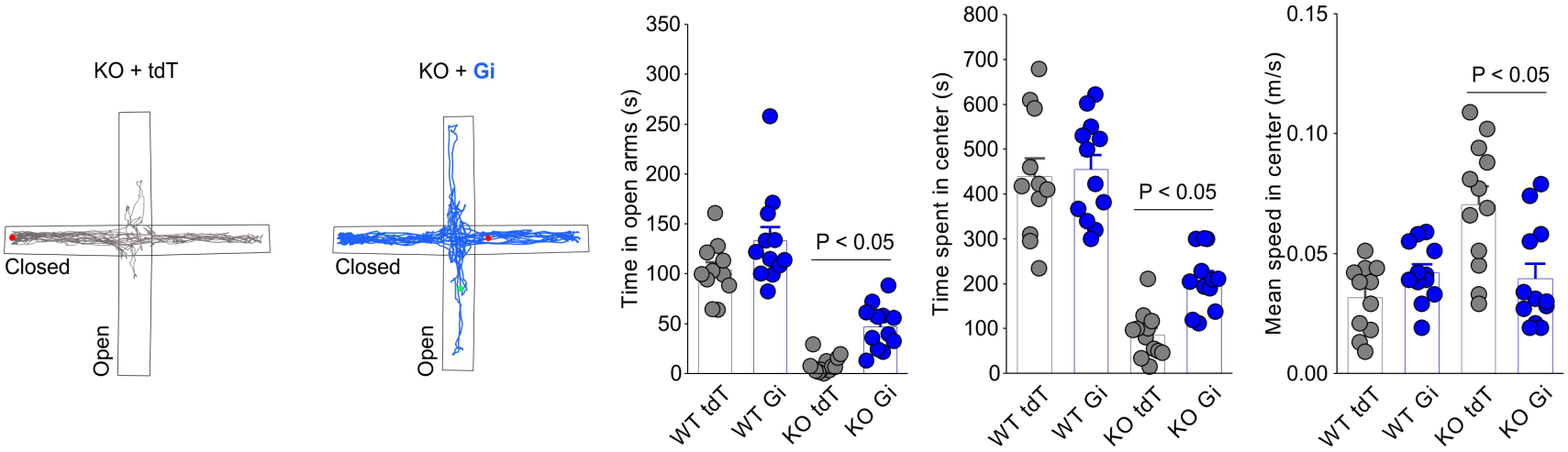

Figure 7.7: Effect of astrocytic G(i) activation on anxiety in SAPAP3 KO mice. Representative traces of 10 minute elevated plus maze recordings for WT and SAPAP3 KO mice treated with DCZ and microinjected with tdTomato or hM4Di Gi GPCR. Green dot depicts the location of the mouse at the start of the recording, while the red dot depicts the location of the mouse at the end of the recording. Bar graphs show anxiety parameters measured by three different metrics: Time spent in open arms of the EPM, mean speed in the center of the OFA, and time spent in the center of the OFA. (n = 11-12 mice per group; Kruskal-Wallis test).

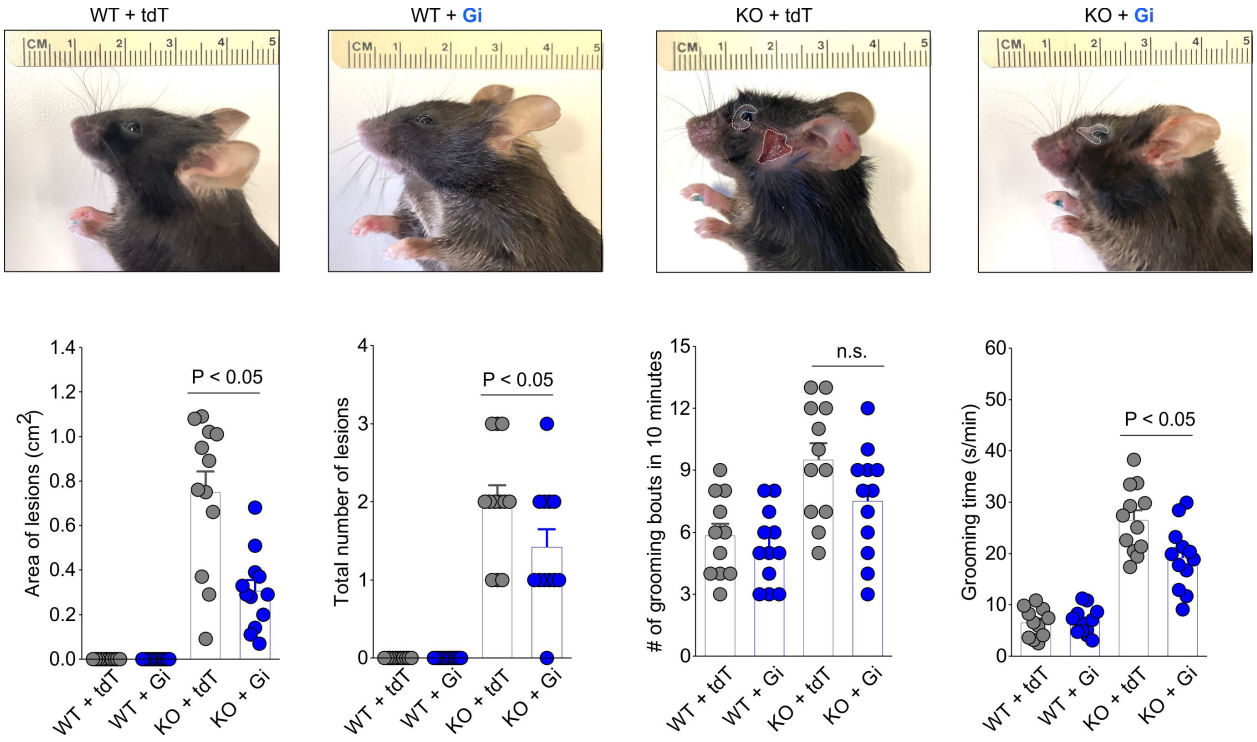

Figure 7.8: Effect of astrocytic G(i) activation on self-grooming in SAPAP3 KO mice. Representative photographs of WT and SAPAP3 KO mice treated with DCZ and microinjected with tdTomato or hM4Di Gi GPCR. White outlines show the extent of the lesions. Bar graphs show self-grooming parameters measured by four different metrics: area of lesions per mouse, total number of lesions per mouse, number of grooming bouts in 10 minutes, and time spent grooming. (n = 11-12 mice per group; Kruskal-Wallis test).

with G(q) GPCR pathway activation, indicating that the G(i) pathway specifically drives improvements via Gi mediated intracellular signaling. These data were supported by our proteomic assessments in which G(i) pathway proteins were shown to interact with SAPAP3 within astrocytes. Thus, the G(i) pathway may be a core astrocytic mechanism involved in OCD-like phenotypes in SAPAP3 KO mice. To further explore this, we assessed cellular readouts of neurons and astrocytes in the context of SAPAP3 KO and subsequent astrocytic Gi activation.

Section 7.4: Cellular correlates of astrocytic Gi activation in SAPAP3 KO mice

Our results indicate that striatal astrocyte G(i) pathway activation can drive changes in repetitive and anxiety behaviors in the SAPAP3 KO context. Our previous cellular data revealed that striatal astrocyte morphology, including the actin cytoskeleton, is altered in SAPAP3 KO mice. Furthermore, by using ΔFosB as a marker of neuronal activity, we found that striatal neurons are hyperactive in SAPAP3 KO mice when compared to WT. Finally, we found that some of these cellular parameters are restored by neuron or astrocyte specific genetic rescue of SAPAP3 (**Chapter 6**). To further assess the role of astrocytic G(i) pathway activation in the context of SAPAP3 KO, we assessed both astrocyte morphology and neuronal hyperactivity in KO mice with activated G(i) pathway in striatal astrocytes.

We evaluated astrocyte morphology and the actin cytoskeleton with an astrocyte-specific LifeAct-GFP AAV (AAV 2/5 *gfaABC$_1$D* LifeAct-GFP*)*. We microinjected wither Gi or astrocyte specific tdTomato in combination with LifeAct GFP bilaterally into the striatum of WT and SAPAP3 KO mice at ~P28. After one month of expression, we began injection of deschloroclozapine (DCZ), a DREADD specific agonist, systemically into all mice intraperitoneally (i.p.) (10μg/kg) for three times a week for three months. There were a total of

four experimental groups: 1) WT + tdTomato, 2) WT + Gi, 3) SAPAP3 KO + tdTomato, and 4) SAPAP3 KO + Gi which were all treated with DCZ. At ~P180, or about 6 months of age, we processed the tissue and conducted immunohistochemistry to assess the astrocyte morphology via LifeAct-GFP signal. When we measured the total actin area of astrocytes across all four groups, we found no difference between the WT + tdT and WT + Gi group. However, consistent with previous analyses, the astrocyte actin area was significantly small in KO + tdT mice but was significantly increased in KO + Gi mice when compared to tdT controls. Analysis of the intensity of actin signal across a 40 μm area supported the LifeAct area measurements as the F-actin intensity at the edges of astrocytes were significantly reduced in the KO + tdT group but this intensity was significantly recovered in the KO + Gi group (**Figure 7.10**). Therefore, activation of the astrocytic G(i) pathway directly restores decreased F-actin and reduced morphology of SAPAP3 KO astrocytes.

We next interrogated broad neuronal activity changes in the striatum using ΔFosB as a marker of sustained neuronal activity. Analysis of the percent of neurons with ΔFosB showed no difference between WT + tdT and WT + Gi groups. As expected, KO + tdT mice showed a significant increase in ΔFosB+ neurons in both the dorsal and ventral striatum when compared to both WT groups, thus confirming neuronal hyperactivity within the striatum. When we analyzed the KO + Gi group, we found that these mice showed significantly reduced ΔFosB+ neurons in both the dorsal and ventral striatum when compared to the KO + tdT group indicating that G(i) pathway activation within astrocytes may drive a specific molecular pathway that modulates neurons to ameliorate repetitive and anxiety behaviors in the context of SAPAP3 KO (**Figure 7.11**). These results are exciting as they provide a molecular basis by which astrocytes contribute to OCD-like phenotypes in mice via SAPAP3.

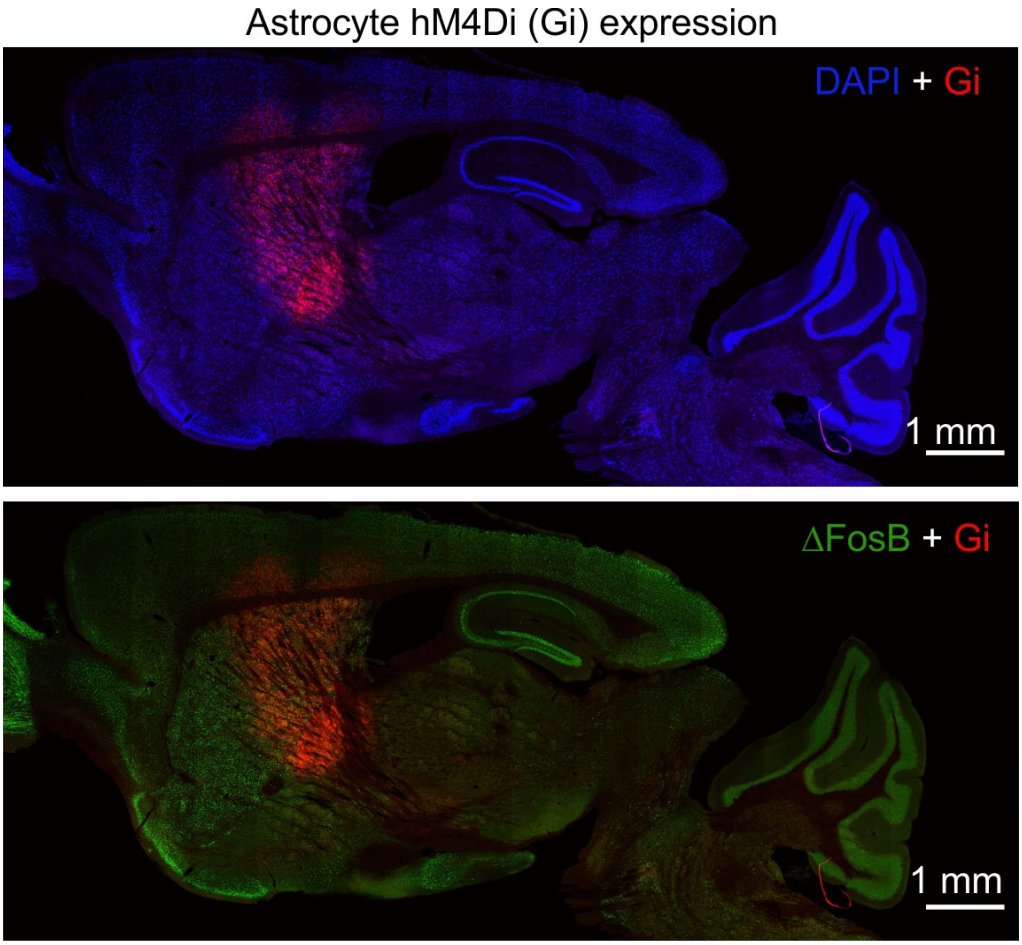

Figure 7.9: Expression of astrocyic hM4Di in the striatum. Representative saggital images showing hM4Di expression (red) in the mouse striatum. Blue, DAPI; green, ΔFosB.

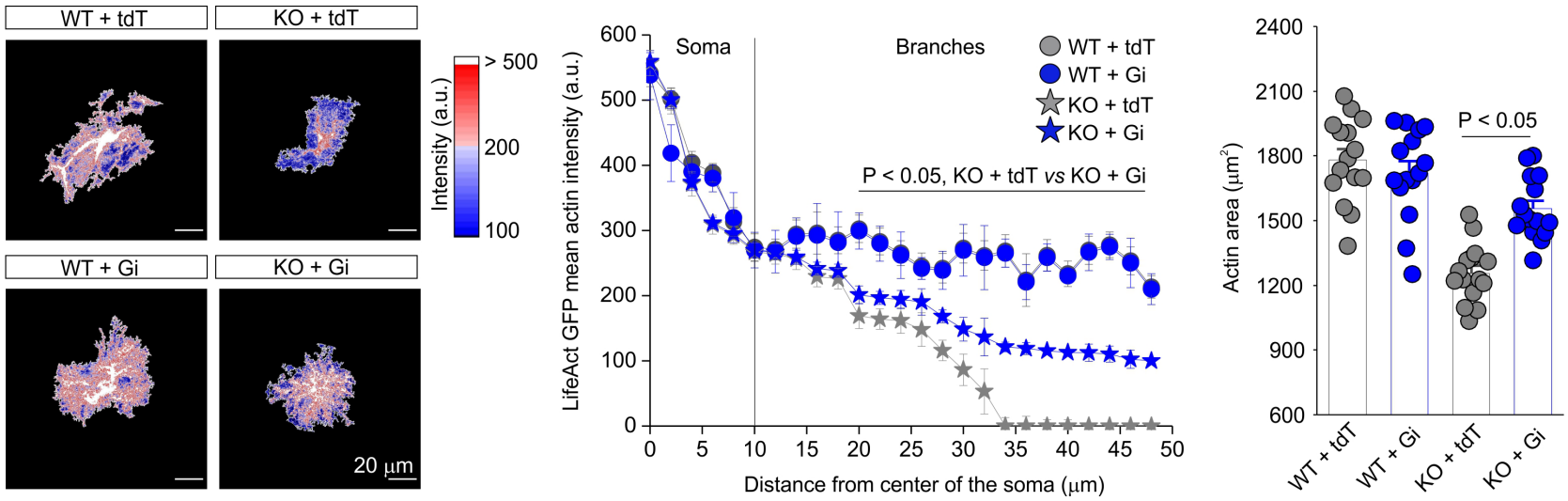

Figure 7.10: Effect of astrocytic G(i) pathway activation on astrocyte morphology. a. Representative images showing LifeAct GFP expression in astrocytes of wild-type and SAPAP3 KO mice treated with DCZ and microinjected with either tdT or hM4Di Gi AAV. Line graph depicts LifeAct GFP intensity as a function of distance from the center of the astrocyte soma. Points represent mean intensity from 14-16 cells per group from 4 mice. (Two-way repeated measures ANOVA with Bonferroni post-hoc test). Bar graphs shows the total LifeAct GFP area in wild-type and SAPAP3 KO mice treated with DCZ and microinjected with either tdT or hM4Di Gi AAV. (n = from 14-16 cells per group from 4 mice; Kruskal-Wallis test)

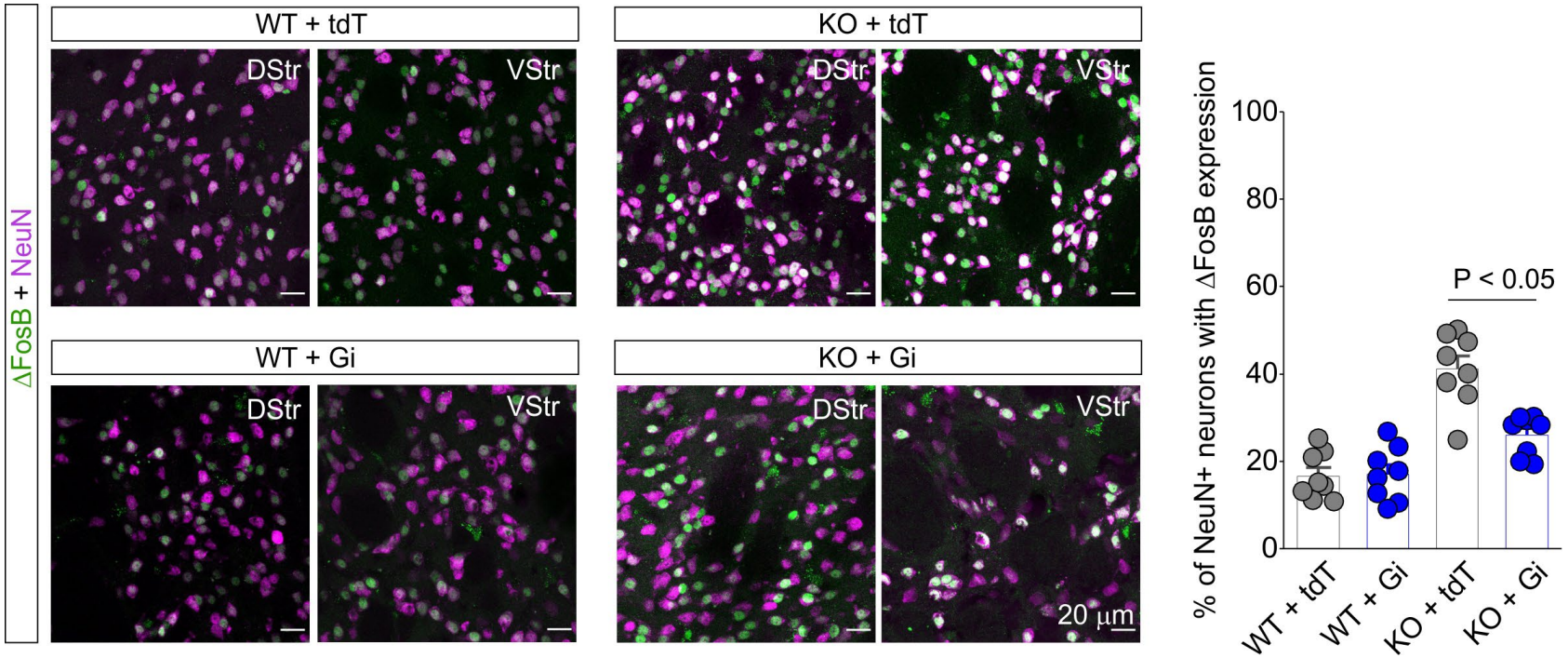

Figure 7.11: Effect of G(i) pathway activation on ΔFosB expression in the striatum. Representative images show ΔFosB expression in dorsal and ventral striatum of wild-type and SAPAP3 KO mice treated with DCZ and microinjected with either tdT or hM4Di Gi AAV. Bar graphs show the percent of neurons with ΔFosB expression in wild-type and SAPAP3 KO mice treated with DCZ and microinjected with either tdT or hM4Di Gi AAV. (n= 4 mice per group; one-way ANOVA with Bonferroni post hoc test).

Section 7.5: Summary

Throughout this dissertation, we assessed the contributions of neurons and astrocytes to repetitive and anxiety behaviors using the SAPAP3 KO mouse model of OCD. We confirmed that fluoxetine, a first line treatment of human OCD, recovers the OCD-like phenotypes in the SAPAP3 KO mouse model. We then discovered that genetic rescue of SAPAP3 in either astrocytes or neurons ameliorates self-grooming and repetitive phenotypes, but have different degrees of contributions in relation to anxiety. Given these results, we finally assessed whether activation of specific GPCR pathways within astrocytes could also modulate OCD-like behaviors in SAPAP3 KO mice. Excitingly, we found that astrocytic activation of the G(i) pathway, but not the G(q) pathway significantly improves both repetitive and anxiety behaviors in SAPAP3 KO mice. To summarize the extent of this improvement and also consolidate all behavioral data presented, we analyzed the data using a Z-score measurement. The Z-score heat map shows the extent rescue for fluoxetine, the astrocyte genetic rescue, the neuronal genetic rescue, and the astrocyte GPCR pathway activation rescue across all nine behavioral parameters we measured (**Figure 7.12**). The heat map underscores that while Gi activation significantly improves behaviors when compared to respective tdT controls, it does not rescue it to the same effect as genetic rescue with Astro SAPAP3. While this may seem like a limitation, we must assess the way we define our goals for possible therapeutic strategies for disorders such as OCD.

OCD remains a poorly treated and understudied psychiatric disease. Presently, there is not one hint for a possible cure, however, treatments such as fluoxetine exist to improve the quality of life for patients. While SSRIs are the first line of treatment for human OCD, only up to 65% of patients respond to fluoxetine and/or other SSRI drugs with about a 25% reduction in symptoms (Koran et al., 2007), thus indicating important limitations with current drug treatments. Therefore,

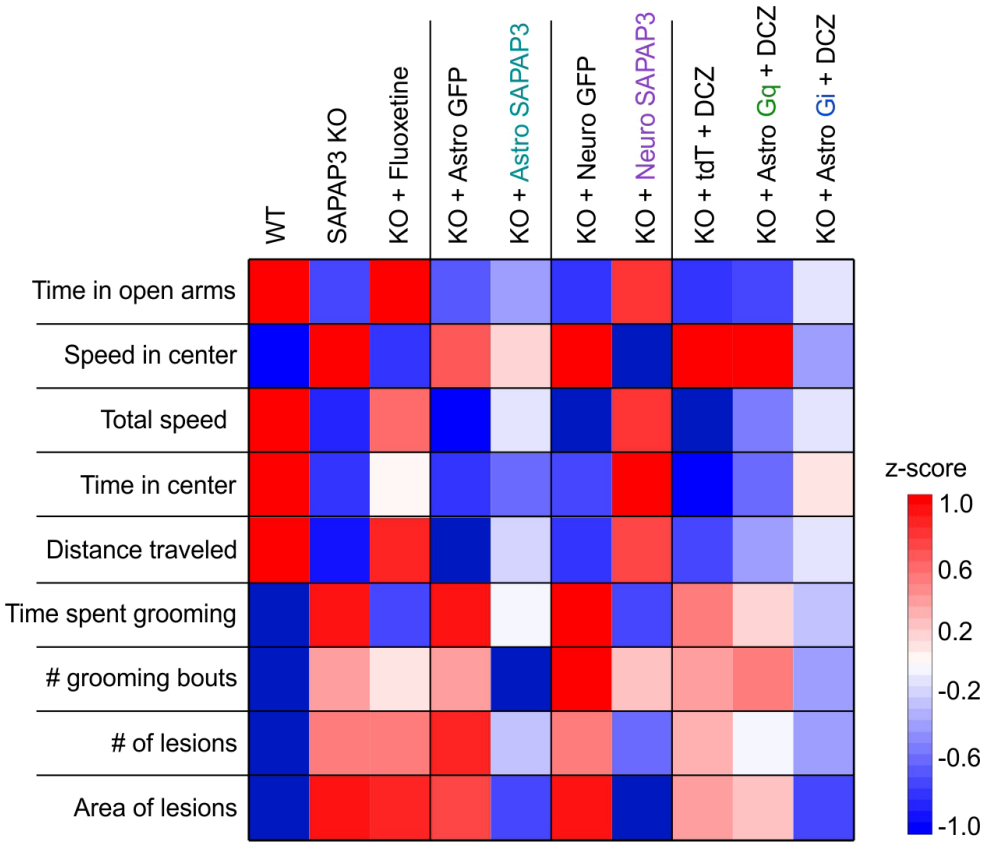

Figure 7.12: Z-score summary of all OCD-related behaviors and their responses to diverse cell-specific interventions. Heat map shows the z-score for all measured OCD-related behaviors across interventions with fluoxetine, cell-specific genetic rescue, and activation of astrocytic G(q) and G(i) G-protein pathways.

identifying and understanding distinct cellular mechanisms related to OCD-phenotypes provide new ideas for novel treatments. In our curiosity to understand basic astrocyte biology, we have found that activating astrocytic G(i) can reduce OCD-like phenotypes in mice and provides a new line of investigation to molecularly dissect pathways that can possibly improve outcomes in OCD. Thus, by understanding basic multicellular interactions, we can probe cell-selective pathways in both astrocytes and neurons that could help us explore novel therapeutic strategies. The exploration of these is an exciting continuation in the quest to understand basic astrocyte function.

CHAPTER 8: CONCLUSIONS

Section 8.1: Data Summary

Since their first description by Virchow in 1858, the role of glia in the nervous system has been the subject of much speculation. This speculation led to novel hypotheses by pioneers such as Golgi and Ramon y Cajal who first documented astrocytes in the brain with their contemporary histology methods. However, the lack of specific tools to study astrocytes caused these cells to become understudied for the majority of the 20th century as neuroscience moved toward a more neuro-centric view in favor of the electrically excitable neurons. With the development of novel cell culturing methods, novel imaging techniques, and the advent of nucleic acid sequencing in the early 21st century, the hypotheses developed by early neuroscientists in relation to astrocytes finally became testable. However, because of their complex morphology and close association with neurons, it is imperative to develop tools to assess astrocyte biology as close to *in vivo* as possible.

Here, we presented a toolset to assess cell and subcompartment-specific proteomes *in vivo* using genetically targeted biotin ligase, BioID2 (**Chapter 3**). With this method we were able to assess the cytosolic and plasma membrane proteomes of neurons and astrocytes in the striatum, a model brain circuit with well-defined inputs and outputs. We validated that our constructs were both cell-selective and functional within neurons and astrocytes and were able to reliably identify proteins which provide a basis for cell-specific signaling. Because genetically targeted BioID2 provided reliable results, we then extended the BioID2 toolset to explore astrocyte subproteomes at five genetically defined functional subcompartments within striatal astrocytes including the astrocyte end foot, fine processes, sites of glutamate buffering, sites of potassium buffering, and astrocyte-astrocyte contact points at gap junctions (**Chapter 4**). The proteins identified at these astrocyte

subcompartments allowed us to define the protein basis of the diverse physiological functions that an astrocyte conducts.

Our cell- and subcompartment- proteomic analyses revealed that both striatal astrocytes and neurons express SAPAP3 (gene: *Dlgap3*), a canonical post-synaptic density protein that is associated with human obsessive-compulsive disorder (OCD). We discovered that SAPAP3 is highly expressed in the astrocyte fine processes and proximal to the plasma membrane where astrocytes interact with neuronal synapses (**Chapters 3, 4, and 5**). This discovery led us to explore the molecular mechanisms of SAPAP3 within astrocytes which included glutamate homeostasis and regulation of the actin cytoskeleton (**Chapter 5**). Furthermore, these molecular functions were altered in SAPAP3 KO mice, which exhibit OCD-like symptoms.

To further define the cell-specific molecular mechanisms of SAPAP3 in the striatum, we genetically rescued SAPAP3 specifically in astrocytes and neurons to assess cell-specific contributions to OCD-related phenotypes in SAPAP3 KO mice (**Chapter 6**). These experiments showed that both astrocytes and neurons contribute to repetitive and compulsion behaviors, but neurons play a larger role in anxiety behaviors. Thus, our data revealed different contributions of astrocytes and neurons to different behaviors related to OCD.

Finally, we explored how modulating distinct signaling pathways within astrocytes could affect OCD-like behaviors in SAPAP3 KO mice (**Chapter 7**). Because our proteomic data revealed that SAPAP3 interacts with G(i) GPCR-pathway signaling molecules and our behavioral data revealed that astrocytes contribute to OCD-behaviors, we hypothesized that activating this pathway could result in some sort of behavioral effect. Our results demonstrated that specific activation of the G(i) GPCR pathway within astrocytes was sufficient to drive behavioral improvements in the SAPAP3 KO mouse model. These results were not replicated with astrocytic G(q) GPCR pathway

activation indicating that the G(i) pathway is specific to this behavioral modulation. Therefore, our results provide a new line of evidence that treatments that could target both astrocytes and neurons should be explored in OCD and perhaps other brain disorders.

In all, the toolset we developed here to explore astrocyte and neuron biology at the protein level provided novel insights into astrocyte function in the striatum and multicellular interactions that drive distinct behavioral phenotypes. All the data presented here are publically available. Our RNA-sequencing data for striatal neurons and astrocytes have been deposited at the Gene Expression Omnibus repository (http://www.ncbi.nlm.nih.gov/geo) with accession number GSE184773. All our neuron and astrocyte proteomics data, including the astrocyte subproteome data have been deposited at the Proteome Exchange Consortium via PRIDE with accession number PXD029257. The plasmids corresponding to the neuron and astrocyte-targeted AAVs developed for this dissertation have been deposited on Addgene.

Section 8.2: Cell- and subcompartment-specific proteomes define the molecular basis of function

Using biotin ligase, BioID2, we were able to reliably capture and identify proteins in both the cytosol and plasma membrane of astrocytes and neurons *in vivo*. Analyses of the proteins identified in both cell types at different compartments revealed that the major signaling pathways in astrocytes were lipid metabolism, cell-cell signaling, and actin filament based processes while the major signaling pathways in neurons were ion binding, receptor signaling, and synaptic signaling. To further probe into these pathways, we evaluated specific proteins within these unique pathways to compare the astrocyte and neuronal proteomes. We found that proteins related to Ca^{2+} dependent vesicular neurotransmitter release were abundant in neurons, but absent in astrocytes (**Chapter 3**).

Furthermore, proteins involved in lipid metabolism were highly abundant in astrocytes with far fewer in neurons. Therefore, our data define a protein basis for cell-type specific signaling mechanisms that diverge the functions conducted by striatal neurons and astrocytes. While seemingly obvious, this functional comparison had not been done and serves to underscore the functions conducted by neurons and astrocytes in the striatum.

Additionally, our data provide a platform to assess and perform hypothesis driven experiments related to cell-type specific function in the striatum. While we did analyze divergent mechanisms in astrocytes and neurons, of interest were also those proteins that were found in both cell types, as they represent a basis for multicellular contributions that drive overall circuit function. We found at least 100 proteins that were shared between astrocytes and neurons that could represent shared mechanisms between both cell types. We validated the expression of ZO-1 (gene: *Tjp1*) a gene canonically described in endothelial cells, and protein Tau (gene: *Mapt*) a protein canonically described in neurons, and found that they were both highly expressed in striatal astrocytes at the protein and RNA-level. This indicates that the way in which a gene or protein is defined as "cell-specific" may have to be re-defined and re-assessed as there could be many proteins that are understudied in distinct cell types simply because of their presumed "cell-specificity." Proteins like these may need to be studied on a case-by-case basis.

A surprising protein shared between astrocytes and neurons was striatum enriched SAPAP3. While traditionally a canonical post-synaptic density protein, we were surprised to see SAPAP3 reliably identified at the astrocyte plasma membrane and within the cytosol indicating it is expressed intracellularly near the plasma membrane of astrocytes. With this discovery, we were able to conduct several hypothesis-driven experiments to understand its function within astrocytes and in relation to obsessive-compulsive disorder. Thus, the cell-type and subcompartment-specific

proteomes provide an unprecedented resource to the scientific community to assess the function of cell-specific proteins within defined neural circuits.

Section 8.3: Subproteomes define molecular functions of complex astrocytes

We studied the proteomes of astrocytes at distinct functional subcompartments because we were inspired by the variety of function served by morphologically complex astrocytes. Because our cell- and subcompartment-specific methods allowed us to reliably detect proteins, we expanded the use of BioID2 to five functional subcompartments. To do this we generated AAVs for Aqp4-BioID2 to assess astrocyte end feet, 2) Ezr-BioID2 to explore astrocyte fine processes, 3) Glt1-BioID2 to assess sites of glutamate uptake, 4) Kir4.1-BioID2 to assess sites of extracellular potassium homeostasis, and 5) Cx43-BioID2 to assess astrocyte-astrocyte contacts. We validated that each construct resulted in BioID2 expression similar to the endogenous protein within astrocytes and also demonstrated functional biotinylation at each subcompartment. This resulted in the discovery of a myriad of subcompartment specific proteins that defined the functions of each subcompartment within astrocytes. Broad assessments of these proteins using gene ontology analyses revealed that the astrocyte subcompartments differed in their predicted biological functions and in their major signaling pathways.

To summarize the entirety of this rich dataset, we created the “astrocyte subproteome cards” which summarize each astrocyte subproteome in a broad and accessible way. For each subproteome card we present 1) unique and enriched proteins, 2) correlations between protein abundance and RNA expression, 3) validation of a candidate protein identified within the targeted subcompartment, 4) a putative protein-protein interaction map for the unique and enriched proteins, and 5) predicted signaling and functional pathways. The astrocyte subproteome cards

reveal known protein-protein interactions as well as hundreds of new putative interactions waiting to be studied. One of these novel interactions was that of Ezrin at the fine processes with SAPAP3, which we extensively studied. Taken together, our astrocyte subcompartment cards provide a platform for design of new experiments in order to understand astrocyte roles in brain function.

Section 8.4: On the relationship between RNA and protein

The advent of RNA-sequencing in the early 21st century provided novel molecular mechanisms by which to study astrocytes. Coupled with genetic strategies to obtain cell-specific RNA, expression analyses using RNA have been invaluable to understand astrocyte roles in health and disease. However, as already discussed, the relationship between RNA expression and protein levels is remarkably complex as post-translational modifications, protein localization mechanisms and RNA-regulation deeply influence the abundance of both macromolecules. Therefore, to reliably assess cellular function, it was imperative for us to assess protein based mechanisms for striatal astrocytes and neurons. While we achieved this goal, we were curious to assess how our proteomic data related to the cell-specific transcriptomic data that our lab has previously described.

To assess this, we compared the RNA abundance from neuron and astrocyte specific RiboTag methods to protein abundance from our neuron and astrocyte specific BioID2 approach. We found that although many of the top astrocyte and neuron enriched genes were also detected in the proteomes, the relationship between the RNA expression and protein abundance was weak. The correlation between RNA and protein was remarkably lower for proteins found at the plasma membrane. Therefore, at least with our methods, protein abundance cannot be accurately predicted by RNA abundance.

We conducted similar analyses across all five astrocyte subproteomes and our data showed that the correlations between RNA and protein were markedly different across all proteomes, but the overarching conclusion was the same: there was very low correlation (pearson's $r < 0.6$ and spearmann's $\rho < 0.5$) between RNA abundance and protein amount within any subproteome. These results are interesting because they may reflect biology related to differences in transcript and protein turnover that could have important implications about a cell's function and requires future exploration.

Section 8.5: Cell-specific contributions in a mouse model of OCD

Using our cell-specific proteomics approaches, we discovered that SAPAP3 is expressed by astrocytes in addition to neurons. This was of interest to us because SAPAP3 is associated with OCD and repetitive behaviors in humans, and these behaviors are primarily driven by the striatum. Thus, we confirmed the expression of SAPAP3 within astrocytes using a variety of –omic approaches including proteomics, deep RNA-seq, single-cell RNA-seq, and histology methods. All the approaches we used confirmed the expression of SAPAP3 within striatal astrocytes. We then evaluated the molecular mechanisms of SAPAP3 function within astrocytes using a range of molecular techniques and discovered that astrocytic SAPAP3 functions to regulate the actin cytoskeleton and possibly, glutamate homeostasis in astrocytes. The actin cytoskeleton regulation function was confirmed with the use of constitutive SAPAP3 KO mice, which revealed decreased territory sizes and altered F-actin within striatal astrocytes. Furthermore, we confirmed known SAPAP3 interactions within neurons including interactions at the glutamatergic post-synaptic density. Given the distinct, but shared molecular mechanisms between neuronal and astrocytic

SAPAP3, we then wanted to assess the relative contributions of both cell types to OCD-related behaviors using the SAPAP3 KO mice.

SAPAP3 KO mice exhibit compulsive self-grooming and elevated anxiety, thus making them an excellent mouse model of OCD. We genetically expressed SAPAP3 in either astrocytes or neurons in the striatum of SAPAP3 KO mice which revealed that both cell types contribute to OCD-phenotypes. Both striatal astrocytes and neurons ameliorate excessive self-grooming upon genetic rescue of SAPAP3, however only neurons were able to rescue the anxiety phenotypes. Thus, this result indicates that OCD-like phenotypes originate and are heavily regulated by the striatum and that multicellular interactions between astrocytes and neurons play important convergent and divergent roles in repetitive and anxiety behaviors. These claims were further supported by our data showing that both astrocyte and neuronal genetic rescue of SAPAP3 were sufficient to lower medium spiny neuron hyperactivity and influenced downstream circuitry differentially. However, only astrocyte specific genetic rescue was sufficient to restored the altered morphology of astrocytes indicating roles for cell-autonomous and non-cell autonomous mechanisms. Our findings thus highlight the molecular, cellular, and behavioral differences and similarities between neuronal and astrocytic mechanisms relevant to OCD-like phenotypes in the SAPAP3 KO mouse model.

Because astrocytes played a causal role in OCD-related behaviors in SAPAP3 KO mice and because our assessments revealed that astrocytic SAPAP3 interacted with proteins involved in the G(i) GPCR pathway, we investigated whether activation of this pathway in striatal astrocytes could modulate OCD-related behaviors. Our studies found that specific activation of the G(i) GPCR pathway, but not orthogonal pathways such as the G(q) GPCR pathway, was sufficient to improve compulsive self-grooming and anxiety in SAPAP3 KO mice. Our results indicate that investigation

of cell-specific signaling mechanisms in astrocytes and neurons may be fruitful to engineer novel therapeutic strategies for OCD or potentially other brain disorders.

Section 8.6: Relevance to human studies

Finally, we explored the implications of our results in relation to human OCD. We conducted bulk proteomics analysis on wild-type and SAPAP3 KO mice to explore differentially expressed proteins. We related our differentially expressed proteins to gene expression changes from human OCD post-mortem tissue and to astrocyte and neuron specific gene expression from our datasets. Our analyses revealed 66 differentially expressed proteins, all of which were expressed to some degree in astrocytes and/or neurons. 65% of the genes were also altered in human OCD, with 27% showing similar directional changes.

Furthermore, we identified the top 30 differentially expressed genes in striatum of human OCD and found that several were also expressed by astrocytes and neurons at the transcriptional level and were found within the proteomes. Finally, we extended our analysis and found 61 genes that were associated with human OCD and Tourette's syndrome from the NIH Phenopedia database. All 61 genes had some degree of expression within our astrocyte and neuron transcriptomes with many found within the astrocyte and neuron proteomes. Therefore, these analyses suggest that indeed, the molecular changes associated with OCD do impact signaling in both astrocytes and neurons. These results also support that understanding basic biological mechanisms in model systems such as mice in a careful and hypothesis-driven manner could reveal important mechanisms relevant to human disorders.

Section 8.7: Conclusion

In all, we were driven by the same curiosity possessed by scientists in the 19th century to understand how astrocytes function in the brain. This motivated us to develop a toolset to assess astrocyte and neuronal proteomes *in vivo*. As the most diverse and complex building blocks in biology, understanding proteins gives us a molecular handle on a cell's function. With the tools we developed, we were able to define the fundamental functional differences between striatal astrocytes and neurons while also defining a molecular basis for the myriad of functions a morphologically complex astrocyte conducts within a defined neural circuit. We have provided a rich and complex dataset that represents previously unexplored interactions and molecules that compliments gene expression data to explore both astrocytes and neurons *in vivo*.

While the relative contributions of astrocytes and neurons to brain disorders has long been speculated, we conducted causal manipulations to assess how astrocytes and neurons molecularly contribute to aberrant phenotypes related to OCD. Our studies revealed that SAPAP3, a protein expressed by both astrocytes and neurons, produces distinct effects on OCD-related behaviors by specific astrocyte and neuron molecular interactions. Our studies show that multicellular interactions are relevant in complex brain diseases and provide a basis for a myriad of novel ideas by which to test these interactions.

CHAPTER 9: BIBLIOGRAPHY

Ahmari, S. E., Spellman, T., Douglass, N. L., Kheirbek, M. A., Simpson, H. B., Deisseroth, K., Gordon, J. A., & Hen, R. (2013). Repeated cortico-striatal stimulation generates persistent OCD-like behavior. *Science*, *340*(6137). https://doi.org/10.1126/science.1234733

Alberts, B., Johnson, A., Lewis, J., Morgan, D., Raff, M., Roberts, K., & Walter, P. (2017). Molecular Biology of the Cell. In *Molecular Biology of the Cell*. https://doi.org/10.1201/9781315735368

Allen, N. J. (2014). Synaptic plasticity: Astrocytes wrap it up. In *Current Biology* (Vol. 24, Issue 15). https://doi.org/10.1016/j.cub.2014.06.030

Allen, N. J., & Lyons, D. A. (2018). System Formation and Function. *Science*, *185*(October).

Alzheimer, A. (1906). Über einen eigenartigen schweren Erkrankungsprozeß der Hirnrinde. *Neurologisches Zentralblatt*, *23*.

Anderson, M. A., Burda, J. E., Ren, Y., Ao, Y., O'Shea, T. M., Kawaguchi, R., Coppola, G., Khakh, B. S., Deming, T. J., & Sofroniew, M. V. (2016). Astrocyte scar formation AIDS central nervous system axon regeneration. *Nature*, *532*(7598). https://doi.org/10.1038/nature17623

Andriezen, W. L. (1894a). On some of the newer aspects of the pathology of insanity: Ambiguous cell-system. *Brain*, *17*(4). https://doi.org/10.1093/brain/17.4.594

Andriezen, W. L. (1894b). On some of the newer aspects of the pathology of insanity: Evolution of polymorphic cell types. *Brain*, *17*(4). https://doi.org/10.1093/brain/17.4.623

Bachoo, R. M., Kim, R. S., Ligon, K. L., Maher, E. A., Brennan, C., Billings, N., Chan, S., Li, C.,

Rowitch, D. H., Wong, W. H., & DePinho, R. A. (2004). Molecular diversity of astrocytes with implications for neurological disorders. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(22). https://doi.org/10.1073/pnas.0402140101

Barker, R. J., Price, R. L., & Gourdie, R. G. (2000). Increased co-localization of Connexin43 and ZO-1 in dissociated adult myocytes. *Cell Adhesion and Communication*, *8*(4–6).

Barres, B. A. (2008). The Mystery and Magic of Glia: A Perspective on Their Roles in Health and Disease. In *Neuron* (Vol. 60, Issue 3). https://doi.org/10.1016/j.neuron.2008.10.013

Barrett, P. J., Song, Y., Van Horn, W. D., Hustedt, E. J., Schafer, J. M., Hadziselimovic, A., Beel, A. J., & Sanders, C. R. (2012). The amyloid precursor protein has a flexible transmembrane domain and binds cholesterol. *Science*, *336*(6085). https://doi.org/10.1126/science.1219988

Bates, G. P., Dorsey, R., Gusella, J. F., Hayden, M. R., Kay, C., Leavitt, B. R., Nance, M., Ross, C. A., Scahill, R. I., Wetzel, R., Wild, E. J., & Tabrizi, S. J. (2015). Huntington disease. In *Nature Reviews Disease Primers* (Vol. 1). https://doi.org/10.1038/nrdp.2015.5

Bayraktar, O. A., Fuentealba, L. C., Alvarez-Buylla, A., & Rowitch, D. H. (2015). Astrocyte development and heterogeneity. *Cold Spring Harbor Perspectives in Biology*, *7*(1). https://doi.org/10.1101/cshperspect.a020362

Bazargani, N., & Attwell, D. (2016). Astrocyte calcium signaling: The third wave. In *Nature Neuroscience* (Vol. 19, Issue 2). https://doi.org/10.1038/nn.4201

Bélanger, M., Allaman, I., & Magistretti, P. J. (2011). Brain energy metabolism: Focus on Astrocyte-neuron metabolic cooperation. In *Cell Metabolism* (Vol. 14, Issue 6). https://doi.org/10.1016/j.cmet.2011.08.016

Benarroch, E. E. (2005). Neuron-astrocyte interactions: Partnership for normal function and disease in the central nervous system. *Mayo Clinic Proceedings*, *80*(10). https://doi.org/10.4065/80.10.1326

Benediktsson, A. M., Schachtele, S. J., Green, S. H., & Dailey, M. E. (2005). Ballistic labeling and dynamic imaging of astrocytes in organotypic hippocampal slice cultures. *Journal of Neuroscience Methods*, *141*(1). https://doi.org/10.1016/j.jneumeth.2004.05.013

Bienvenu, O. J., Wang, Y., Shugart, Y. Y., Welch, J. M., Grados, M. A., Fyer, A. J., Rauch, S. L., McCracken, J. T., Rasmussen, S. A., Murphy, D. L., Cullen, B., Valle, D., Hoehn-Saric, R., Greenberg, B. D., Pinto, A., Knowles, J. A., Piacentini, J., Pauls, D. L., Liang, K. Y., … Nestadt, G. (2009). Sapap3 and pathological grooming in humans: Results from the OCD collaborative genetics study. *American Journal of Medical Genetics, Part B: Neuropsychiatric Genetics*, *150*(5). https://doi.org/10.1002/ajmg.b.30897

Bordey, A., & Sontheimer, H. (1999). Differential inhibition of glial K(+) currents by 4-AP. *Journal of Neurophysiology*, *82*(6), 3476–3487. https://doi.org/10.1152/jn.1999.82.6.3476

Burguière, E., Monteiro, P., Feng, G., & Graybiel, A. M. (2013). Optogenetic stimulation of lateral orbitofronto-striatal pathway suppresses compulsive behaviors. *Science*, *340*(6137). https://doi.org/10.1126/science.1232380

Burguière, E., Monteiro, P., Mallet, L., Feng, G., & Graybiel, A. M. (2015). Striatal circuits, habits, and implications for obsessive-compulsive disorder. In *Current Opinion in Neurobiology* (Vol. 30). https://doi.org/10.1016/j.conb.2014.08.008

Bushong, E. A., Martone, M. E., Jones, Y. Z., & Ellisman, M. H. (2002). Protoplasmic astrocytes in CA1 stratum radiatum occupy separate anatomical domains. *Journal of Neuroscience*, *22*(1). https://doi.org/10.1523/jneurosci.22-01-00183.2002

Cahoy, J. D., Emery, B., Kaushal, A., Foo, L. C., Zamanian, J. L., Christopherson, K. S., Xing, Y., Lubischer, J. L., Krieg, P. A., Krupenko, S. A., Thompson, W. J., & Barres, B. A. (2008). A

transcriptome database for astrocytes, neurons, and oligodendrocytes: A new resource for understanding brain development and function. *Journal of Neuroscience*, *28*(1). https://doi.org/10.1523/JNEUROSCI.4178-07.2008

Cestra, G., Toomre, D., Chang, S., & De Camilli, P. (2005). The Abl/Arg substrate ArgBP2/nArgBP2 coordinates the function of multiple regulatory mechanisms converging on the actin cytoskeleton. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(5). https://doi.org/10.1073/pnas.0409376102

Chai, H., Diaz-castro, B., Shigetomi, E., Coppola, G., Khakh, B. S., Whitelegge, J. P., Coppola, G., & Khakh, B. S. (2017). Neural Circuit-Specialized Astrocytes : NeuroResource Neural Circuit-Specialized Astrocytes : *Neuron*, *95*.

Charles, A. C., Merrill, J. E., Dirksen, E. R., & Sandersont, M. J. (1991). Intercellular signaling in glial cells: Calcium waves and oscillations in response to mechanical stimulation and glutamate. *Neuron*, *6*(6). https://doi.org/10.1016/0896-6273(91)90238-U

Chen, L., Chetkovich, D. M., Petralia, R. S., Sweeney, N. T., Kawasaki, Y., Wenthold, R. J., Bredt, D. S., & Nicoll, R. A. (2000). Stargazin regulates synaptic targeting of AMPA receptors by two distinct mechanisms. *Nature*, *408*(6815). https://doi.org/10.1038/35050030

Chen, S. K., Tvrdik, P., Peden, E., Cho, S., Wu, S., Spangrude, G., & Capecchi, M. R. (2010). Hematopoietic origin of pathological grooming in Hoxb8 mutant mice. *Cell*, *141*(5). https://doi.org/10.1016/j.cell.2010.03.055

Choi, H., Liu, G., Mellacheruvu, D., Tyers, M., Gingras, A. C., & Nesvizhskii, A. I. (2012). Analyzing protein-protein interactions from affinity purification-mass spectrometry data with SAINT. *Current Protocols in Bioinformatics*, *SUPPL.39*. https://doi.org/10.1002/0471250953.bi0815s39

Clarke, L. E., Liddelow, S. A., Chakraborty, C., Münch, A. E., Heiman, M., & Barres, B. A. (2018). Normal aging induces A1-like astrocyte reactivity. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(8). https://doi.org/10.1073/pnas.1800165115

Corada, M., Mariotti, M., Thurston, G., Smith, K., Kunkel, R., Brockhaus, M., Lampugnani, M. G., Martin-Padura, I., Stoppacciaro, A., Ruco, L., Mcdonald, D. M., Ward, P. A., & Dejana, E. (1999). Vascular endothelial-cadherin is an important determinant of microvascular integrity in vivo. *Proceedings of the National Academy of Sciences of the United States of America*, *96*(17). https://doi.org/10.1073/pnas.96.17.9815

Cox, J., & Mann, M. (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nature Biotechnology*, *26*(12). https://doi.org/10.1038/nbt.1511

Danbolt, N. C. (2001). Glutamate uptake. In *Progress in Neurobiology* (Vol. 65, Issue 1). https://doi.org/10.1016/S0301-0082(00)00067-8

Dani, J. W., Chernjavsky, A., & Smith, S. J. (1992). Neuronal activity triggers calcium waves in hippocampal astrocyte networks. *Neuron*, *8*(3). https://doi.org/10.1016/0896-6273(92)90271-E

Derouiche, A., & Geiger, K. D. (2019). Perspectives for ezrin and radixin in astrocytes: Kinases, functions and pathology. In *International Journal of Molecular Sciences* (Vol. 20, Issue 15). https://doi.org/10.3390/ijms20153776

Dierig, S. (1994). Extending the neuron doctrine: Carl Ludwig Schleich (1859-1922) and his reflections on neuroglia at the inception of the neural-network concept in 1894. *Trends in Neurosciences*, *17*(11). https://doi.org/10.1016/0166-2236(94)90129-5

Endo, F., Kasai, A., Soto, J. S., Yu, X., Qu, Z., Hashimoto, H., Gradinaru, V., Kawaguchi, R., & Khakh, B. S. (2022). Molecular basis of astrocyte diversity and morphology across the CNS in health and disease. *Science*, *378*(6619). https://doi.org/10.1126/science.adc9020

Filipic, B., Gradisnik, L., & Mazija, H. (2022). The astrocytes and astroglia morphology and their biological properties. In *Astrocytes and Their Role in Health and Disease*.

Gangwani, M. R., Soto, J. S., Jami-Alahmadi, Y., Tiwari, S., Kawaguchi, R., Wohlschlegel, J. A., & Khakh, B. S. (2023). Neuronal and astrocytic contributions to Huntington's disease dissected with zinc finger protein transcriptional repressors. *Cell Reports*, *42*(1). https://doi.org/10.1016/j.celrep.2022.111953

Giepmans, B. N. G., & Moolenaar, W. H. (1998). The gap junction protein connexin43 interacts with the second PDZ domain of the zona occludens-1 protein. *Current Biology*, *8*(16). https://doi.org/10.1016/S0960-9822(07)00375-2

Graybiel, A. M. (2008). Habits, rituals, and the evaluative brain. In *Annual Review of Neuroscience* (Vol. 31). https://doi.org/10.1146/annurev.neuro.29.051605.112851

Graybiel, A. M., & Grafton, S. T. (2015). The striatum: Where skills and habits meet. *Cold Spring Harbor Perspectives in Biology*, *7*(8). https://doi.org/10.1101/cshperspect.a021691

Gremel, C. M., & Costa, R. M. (2013). Orbitofrontal and striatal circuits dynamically encode the shift between goal-directed and habitual actions. *Nature Communications*, *4*. https://doi.org/10.1038/ncomms3264

Halassa, M. M., & Haydon, P. G. (2009). Integrated brain circuits: Astrocytic networks modulate neuronal activity and behavior. In *Annual Review of Physiology* (Vol. 72). https://doi.org/10.1146/annurev-physiol-021909-135843

Han, D., Jin, J., Woo, J., Min, H., & Kim, Y. (2014). Proteomic analysis of mouse astrocytes and their secretome by a combination of FASP and StageTip-based, high pH, reversed-phase fractionation. *Proteomics*, *14*(13–14). https://doi.org/10.1002/pmic.201300495

Han, X., Aslanian, A., & Yates, J. R. (2008). Mass spectrometry for proteomics. In *Current Opinion in Chemical Biology* (Vol. 12, Issue 5). https://doi.org/10.1016/j.cbpa.2008.07.024

Heiman, M., Schaefer, A., Gong, S., Peterson, J. D., Day, M., Ramsey, K. E., Suárez-Fariñas, M., Schwarz, C., Stephan, D. A., Surmeier, D. J., Greengard, P., & Heintz, N. (2008). A Translational Profiling Approach for the Molecular Characterization of CNS Cell Types. *Cell*, *135*(4). https://doi.org/10.1016/j.cell.2008.10.028

Hibino, H., Horio, Y., Inanobe, A., Doi, K., Ito, M., Yamada, M., Gotow, T., Uchiyama, Y., Kawamura, M., Kubo, T., & Kurachi, Y. (1997). An ATP-dependent inwardly rectifying potassium channel, KAB-2 (Kir4. 1), in cochlear stria vascularis of inner ear: its specific subcellular localization and correlation with the formation of endocochlear potential. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *17*(12), 4711–4721. https://doi.org/10.1523/JNEUROSCI.17-12-04711.1997

Hosp, F., & Mann, M. (2017). A Primer on Concepts and Applications of Proteomics in Neuroscience. In *Neuron* (Vol. 96, Issue 3). https://doi.org/10.1016/j.neuron.2017.09.025

Jami-Alahmadi, Y., Pandey, V., Mayank, A. K., & Wohlschlegel, J. A. (2021). A robust method for packing high resolution c18 RP-nano-HPLC columns. *Journal of Visualized Experiments*, *2021*(171). https://doi.org/10.3791/62380

Jiang, R., Diaz-Castro, B., Looger, L. L., & Khakh, B. S. (2016). Dysfunctional calcium and glutamate signaling in striatal astrocytes from Huntington's disease model mice. *Journal of Neuroscience*, *36*(12). https://doi.org/10.1523/JNEUROSCI.3693-15.2016

John Lin, C. C., Yu, K., Hatcher, A., Huang, T. W., Lee, H. K., Carlson, J., Weston, M. C., Chen, F., Zhang, Y., Zhu, W., Mohila, C. A., Ahmed, N., Patel, A. J., Arenkiel, B. R., Noebels, J. L., Creighton, C. J., & Deneen, B. (2017). Identification of diverse astrocyte populations and their malignant analogs. *Nature Neuroscience*, *20*(3). https://doi.org/10.1038/nn.4493

Kalueff, A. V., Wayne Aldridge, J., Laporte, J. L., Murphy, D. L., & Tuohimaa, P. (2007). Analyzing grooming microstructure in neurobehavioral experiments. *Nature Protocols*, *2*(10). https://doi.org/10.1038/nprot.2007.367

Kang, J., Jiang, L., Goldman, S. A., & Nedergaard, M. (1998). Astrocyte-mediated potentiation of inhibitory synaptic transmission. *Nature Neuroscience*, *1*(8). https://doi.org/10.1038/3684

Kawaguchi, Y., Wilson, C. J., Augood, S. J., & Emson, P. C. (1995). Striatal interneurones: chemical, physiological and morphological characterization. In *Trends in Neurosciences* (Vol. 18, Issue 12). https://doi.org/10.1016/0166-2236(95)98374-8

Khakh, B. S. (2019). Astrocyte–Neuron Interactions in the Striatum: Insights on Identity, Form, and Function. In *Trends in Neurosciences* (Vol. 42, Issue 9). https://doi.org/10.1016/j.tins.2019.06.003

Kim, D. I., Jensen, S. C., Noble, K. A., Kc, B., Roux, K. H., Motamedchaboki, K., & Roux, K. J. (2016). An improved smaller biotin ligase for BioID proximity labeling. *Molecular Biology of the Cell*, *27*(8). https://doi.org/10.1091/mbc.E15-12-0844

Kim, E., Naisbitt, S., Hsueh, Y. P., Rao, A., Rothschild, A., Craig, A. M., & Sheng, M. (1997). GKAP, a novel synaptic protein that interacts with the guanylate kinase- like domain of the PSD-95/SAP90 family of channel clustering molecules. *Journal of Cell Biology*, *136*(3). https://doi.org/10.1083/jcb.136.3.669

Kim, E., & Sheng, M. (2004). PDZ domain proteins of synapses. In *Nature Reviews Neuroscience* (Vol. 5, Issue 10). https://doi.org/10.1038/nrn1517

Kofuji, P., & Newman, E. A. (2004). Potassium buffering in the central nervous system. *Neuroscience*, *129*(4). https://doi.org/10.1016/j.neuroscience.2004.06.008

Koran, L. M., Hanna, G. L., Hollander, E., Nestadt, G., & Simpson, H. B. (2007). Practice guideline for the treatment of patients with obsessive-compulsive disorder. In *The American journal of psychiatry* (Vol. 164, Issue 7 Suppl). https://doi.org/10.1176/appi.books.9780890423363.149114

Kreitzer, A. C. (2009). Physiology and pharmacology of striatal neurons. In *Annual Review of Neuroscience* (Vol. 32). https://doi.org/10.1146/annurev.neuro.051508.135422

Kreitzer, A. C., & Malenka, R. C. (2008). Striatal Plasticity and Basal Ganglia Circuit Function. In *Neuron* (Vol. 60, Issue 4). https://doi.org/10.1016/j.neuron.2008.11.005

Kügler, S., Kilic, E., & Bähr, M. (2003). Human synapsin 1 gene promoter confers highly neuron-specific long-term transgene expression from an adenoviral vector in the adult rat brain depending on the transduced area. In *Gene Therapy* (Vol. 10, Issue 4). https://doi.org/10.1038/sj.gt.3301905

Kwon, J. S., Jang, J. H., Choi, J. S., & Kang, D. H. (2009). Neuroimaging in obsessive-compulsive disorder. In *Expert Review of Neurotherapeutics* (Vol. 9, Issue 2). https://doi.org/10.1586/14737175.9.2.255

Lanciego, J. L., Luquin, N., & Obeso, J. A. (2012). Functional neuroanatomy of the basal ganglia. *Cold Spring Harbor Perspectives in Medicine*, *2*(12). https://doi.org/10.1101/cshperspect.a009621

Larance, M., & Lamond, A. I. (2015). Multidimensional proteomics for cell biology. In *Nature Reviews Molecular Cell Biology* (Vol. 16, Issue 5). https://doi.org/10.1038/nrm3970

Lee, K. F. H., Soares, C., Thivierge, J. P., & Béïque, J. C. (2016). Correlated Synaptic Inputs Drive Dendritic Calcium Amplification and Cooperative Plasticity during Clustered Synapse Development. *Neuron*, *89*(4). https://doi.org/10.1016/j.neuron.2016.01.012

Lee, Y., Messing, A., Su, M., & Brenner, M. (2008). GFAP promoter elements required for region-specific and astrocyte-specific expression. *GLIA*, *56*(5). https://doi.org/10.1002/glia.20622

Lehre, K. P., & Danbolt, N. C. (1998). The number of glutamate transport subtype molecules at glutamatergic synapses: Chemical and stereological quantification in young adult rat brain. *Journal of Neuroscience*, *18*(21). https://doi.org/10.1523/jneurosci.18-21-08751.1998

Liu, Y., Beyer, A., & Aebersold, R. (2016). On the Dependency of Cellular Protein Levels on mRNA Abundance. In *Cell* (Vol. 165, Issue 3). https://doi.org/10.1016/j.cell.2016.03.014

Lundgaard, I., Osório, M. J., Kress, B. T., Sanggaard, S., & Nedergaard, M. (2014). White matter astrocytes in health and disease. In *Neuroscience* (Vol. 276). https://doi.org/10.1016/j.neuroscience.2013.10.050

Mader, S., & Brimberg, L. (2019). Aquaporin-4 water channel in the brain and its implication for health and disease. In *Cells* (Vol. 8, Issue 2). https://doi.org/10.3390/cells8020090

Maia, T. V., Cooney, R. E., & Peterson, B. S. (2008). The neural bases of obsessive - Compulsive disorder in children and adults. In *Development and Psychopathology* (Vol. 20, Issue 4). https://doi.org/10.1017/S0954579408000606

Marguerat, S., Schmidt, A., Codlin, S., Chen, W., Aebersold, R., & Bähler, J. (2012). Quantitative analysis of fission yeast transcriptomes and proteomes in proliferating and quiescent cells. *Cell*, *151*(3). https://doi.org/10.1016/j.cell.2012.09.019

Martínez-Villarreal, J., García Tardón, N., Ibáñez, I., Giménez, C., & Zafra, F. (2012). Cell surface turnover of the glutamate transporter GLT-1 is mediated by ubiquitination/deubiquitination. *GLIA*, *60*(9). https://doi.org/10.1002/glia.22354

Moye, S. L., Diaz-Castro, B., Gangwani, M. R., & Khakh, B. S. (2019). Visualizing Astrocyte Morphology Using Lucifer Yellow Iontophoresis. *Journal of Visualized Experiments*, *151*. https://doi.org/10.3791/60225-v

Nadler, J. (2012). Plasticity of Glutamate Synaptic Mechanisms Jasper ' s Basic Mechanisms of the Epilepsies. *Jasper ' s Basic Mechanisms of the Epilepsies*.

Nagai, J., Rajbhandari, A. K., Gangwani, M. R., Hachisuka, A., Coppola, G., Masmanidis, S. C., Fanselow, M. S., & Khakh, B. S. (2019). Hyperactivity with Disrupted Attention by Activation of an Astrocyte Synaptogenic Cue. *Cell*, *177*(5), 1280-1292.e20. https://doi.org/10.1016/j.cell.2019.03.019

Nagai, J., Yu, X., Papouin, T., Cheong, E., Freeman, M. R., Monk, K. R., Hastings, M. H., Haydon, P. G., Rowitch, D., Shaham, S., & Khakh, B. S. (2021). Behaviorally consequential astrocytic regulation of neural circuits. In *Neuron* (Vol. 109, Issue 4). https://doi.org/10.1016/j.neuron.2020.12.008

Naisbitt, S., Eunjoon, K., Tu, J. C., Xiao, B., Sala, C., Valtschanoff, J., Weinberg, R. J., Worley, P. F., & Sheng, M. (1999). Shank, a novel family of postsynaptic density proteins that binds to the NMDA receptor/PSD-95/GKAP complex and cortactin. *Neuron*, *23*(3). https://doi.org/10.1016/S0896-6273(00)80809-0

Nicola, S. M., & Deadwyler, S. A. (2000). Firing rate of nucleus accumbens neurons is dopamine-dependent and reflects the timing of cocaine-seeking behavior in rats on a progressive ratio schedule of reinforcement. *Journal of Neuroscience*, *20*(14). https://doi.org/10.1523/jneurosci.20-14-05526.2000

Nielsen, S., Nagelhus, E. A., Amiry-Moghaddam, M., Bourque, C., Agre, P., & Ottersen, O. R. (1997). Specialized membrane domains for water transport in glial cells: High- resolution

immunogold cytochemistry of aquaporin-4 in rat brain. *Journal of Neuroscience*, *17*(1). https://doi.org/10.1523/jneurosci.17-01-00171.1997

Nuriya, M., & Yasui, M. (2013). Endfeet serve as diffusion-limited subcellular compartments in astrocytes. *Journal of Neuroscience*, *33*(8). https://doi.org/10.1523/JNEUROSCI.3050-12.2013

Octeau, J. C., Chai, H., Jiang, R., Bonanno, S. L., Martin, K. C., & Khakh, B. S. (2018). An Optical Neuron-Astrocyte Proximity Assay at Synaptic Distance Scales. *Neuron*, *98*(1). https://doi.org/10.1016/j.neuron.2018.03.003

Ogata, K., & Kosaka, T. (2002). Structural and quantitative analysis of astrocytes in the mouse hippocampus. *Neuroscience*, *113*(1). https://doi.org/10.1016/S0306-4522(02)00041-6

O'Kane, R. L., Martínez-López, I., DeJoseph, M. R., Viña, J. R., & Hawkins, R. A. (1999). Na+-dependent glutamate transporters (EAAT1, EAAT2, and EAAT3) of the blood-brain barrier. A mechanism for glutamate removal. *Journal of Biological Chemistry*, *274*(45). https://doi.org/10.1074/jbc.274.45.31891

Oliet, S. H. R., Piet, R., & Poulain, D. A. (2001). Control of glutamate clearance and synaptic efficacy by glial coverage of neurons. *Science*, *292*(5518). https://doi.org/10.1126/science.1059162

Paukert, M., Agarwal, A., Cha, J., Doze, V. A., Kang, J. U., & Bergles, D. E. (2014). Norepinephrine controls astroglial responsiveness to local circuit activity. *Neuron*, *82*(6). https://doi.org/10.1016/j.neuron.2014.04.038

Penes, M. C., Li, X., & Nagy, J. I. (2005). Expression of zonula occludens-1 (ZO-1) and the transcription factor ZO-1-associated nucleic acid-binding protein (ZONAB)-MsY3 in glial cells and colocalization at oligodendrocyte and astrocyte gap junctions in mouse brain. *European Journal of Neuroscience*, *22*(2). https://doi.org/10.1111/j.1460-9568.2005.04225.x

Rakers, C., Schleif, M., Blank, N., Matušková, H., Ulas, T., Händler, K., Torres, S. V., Schumacher, T., Tai, K., Schultze, J. L., Jackson, W. S., & Petzold, G. C. (2019). Stroke target identification guided by astrocyte transcriptome analysis. *GLIA*, *67*(4). https://doi.org/10.1002/glia.23544

Rasmussen, A. H., Rasmussen, H. B., & Silahtaroglu, A. (2017). The DLGAP family: Neuronal expression, function and role in brain disorders. In *Molecular Brain* (Vol. 10, Issue 1). https://doi.org/10.1186/s13041-017-0324-9

Rauch, S. L., Savage, C. R., Alpert, N. M., Dougherty, D., Kendrick, A., Curran, T., Brown, H. D., Manzo, P., Fischman, A. J., & Jenike, M. A. (1997). Probing striatal function in obsessive-compulsive disorder: A PET study of implicit sequence learning. *Journal of Neuropsychiatry and Clinical Neurosciences*, *9*(4). https://doi.org/10.1176/jnp.9.4.568

Rosenbaum, D. M., Rasmussen, S. G. F., & Kobilka, B. K. (2009). The structure and function of G-protein-coupled receptors. In *Nature* (Vol. 459, Issue 7245). https://doi.org/10.1038/nature08144

Rotge, J. Y., Guehl, D., Dilharreguy, B., Cuny, E., Tignol, J., Bioulac, B., Allard, M., Burbaud, P., & Aouizerate, B. (2008). Provocation of obsessive-compulsive symptoms: A quantitative voxel-based meta-analsysis of functional neuroimaging studies. In *Journal of Psychiatry and Neuroscience* (Vol. 33, Issue 5).

Roth, B. L. (2016). DREADDs for Neuroscientists. In *Neuron* (Vol. 89, Issue 4). https://doi.org/10.1016/j.neuron.2016.01.040

Roux, K. J., Kim, D. I., Burke, B., & May, D. G. (2018). BioID: A Screen for Protein-Protein Interactions. *Current Protocols in Protein Science*, *91*(1). https://doi.org/10.1002/cpps.51

Salmon, C. K., Syed, T. A., Kacerovsky, J. B., Alivodej, N., Schober, A. L., Sloan, T. F. W., Pratte, M. T., Rosen, M. P., Green, M., Chirgwin-Dasgupta, A., Mehta, S., Jilani, A., Wang, Y., Vali, H., Mandato, C. A., Siddiqi, K., & Murai, K. K. (2023). Organizing principles of astrocytic nanoarchitecture in the mouse cerebral cortex. *Current Biology*, *33*(5). https://doi.org/10.1016/j.cub.2023.01.043

Saunders, A., Macosko, E. Z., Wysoker, A., Goldman, M., Krienen, F. M., de Rivera, H., Bien, E., Baum, M., Bortolin, L., Wang, S., Goeva, A., Nemesh, J., Kamitaki, N., Brumbaugh, S., Kulp, D., & McCarroll, S. A. (2018). Molecular Diversity and Specializations among the Cells of the Adult Mouse Brain. *Cell*, *174*(4). https://doi.org/10.1016/j.cell.2018.07.028

Saxena, S., & Rauch, S. L. (2000). Functional neuroimaging and the neuroanatomy of obsessive-compulsive disorder. *Psychiatric Clinics of North America*, *23*(3). https://doi.org/10.1016/S0193-953X(05)70181-7

Scannevin, R. H., & Huganir, R. L. (2000). Postsynaptic organisation and regulation of excitatory synapses. *Nature Reviews Neuroscience*, *1*(2). https://doi.org/10.1038/35039075

Sears, R. M., May, D. G., & Roux, K. J. (2019). BioID as a tool for protein-proximity labeling in living cells. In *Methods in Molecular Biology* (Vol. 2012). https://doi.org/10.1007/978-1-4939-9546-2_15

Sheng, M., & Hoogenraad, C. C. (2007). The postsynaptic architecture of excitatory synapses: A more quantitative view. In *Annual Review of Biochemistry* (Vol. 76). https://doi.org/10.1146/annurev.biochem.76.060805.160029

Sheng, M., & Kim, E. (2011). The postsynaptic organization of synapses. *Cold Spring Harbor Perspectives in Biology*, *3*(12). https://doi.org/10.1101/cshperspect.a005678

Shigetomi, E., Bushong, E. A., Haustein, M. D., Tong, X., Jackson-Weaver, O., Kracun, S., Xu, J., Sofroniew, M. V., Ellisman, M. H., & Khakh, B. S. (2013). Imaging calcium

microdomains within entire astrocyte territories and endfeet with GCaMPs expressed using adeno-associated viruses. *Journal of General Physiology*, *141*(5). https://doi.org/10.1085/jgp.201210949

Shigetomi, E., Patel, S., & Khakh, B. S. (2016). Probing the Complexities of Astrocyte Calcium Signaling. In *Trends in Cell Biology* (Vol. 26, Issue 4). https://doi.org/10.1016/j.tcb.2016.01.003

Smith, S. J. (1994). Neural Signalling: Neuromodulatory astrocytes. *Current Biology*, *4*(9). https://doi.org/10.1016/S0960-9822(00)00178-0

Spray, D. C. (1998). Gap junction proteins: Where they live and how they die. In *Circulation Research* (Vol. 83, Issue 6). https://doi.org/10.1161/01.RES.83.6.679

Srinivasan, R., Huang, B. S., Venugopal, S., Johnston, A. D., Chai, H., Zeng, H., Golshani, P., & Khakh, B. S. (2015). Ca2+ signaling in astrocytes from Ip3r2 -/- mice in brain slices and during startle responses in vivo. *Nature Neuroscience*, *18*(5). https://doi.org/10.1038/nn.4001

Srinivasan, R., Lu, T. Y., Chai, H., Xu, J., Huang, B. S., Golshani, P., Coppola, G., & Khakh, B. S. (2016). New Transgenic Mouse Lines for Selectively Targeting Astrocytes and Studying Calcium Signals in Astrocyte Processes In Situ and In Vivo. *Neuron*, *92*(6). https://doi.org/10.1016/j.neuron.2016.11.030

Stout, R. F., Snapp, E. L., & Spray, D. C. (2015). Connexin type and fluorescent protein fusion tag determine structural stability of gap junction plaques. *Journal of Biological Chemistry*, *290*(39). https://doi.org/10.1074/jbc.M115.659979

Takeuchi, M., Hata, Y., Hirao, K., Toyoda, A., Irie, M., & Takai, Y. (1997). SAPAPs. A family of PSD-95/SAP90-associated proteins localized at postsynaptic density. *Journal of Biological Chemistry*, *272*(18). https://doi.org/10.1074/jbc.272.18.11943

Tang, Y. C., & Amon, A. (2013). Gene copy-number alterations: A cost-benefit analysis. In *Cell* (Vol. 152, Issue 3). https://doi.org/10.1016/j.cell.2012.11.043

Tian, R., Gachechiladze, M. A., Ludwig, C. H., Laurie, M. T., Hong, J. Y., Nathaniel, D., Prabhu, A. V., Fernandopulle, M. S., Patel, R., Abshari, M., Ward, M. E., & Kampmann, M. (2019). CRISPR Interference-Based Platform for Multimodal Genetic Screens in Human iPSC-Derived Neurons. *Neuron*, *104*(2). https://doi.org/10.1016/j.neuron.2019.07.014

Trapnell, C., Williams, B. A., Pertea, G., Mortazavi, A., Kwan, G., Van Baren, M. J., Salzberg, S. L., Wold, B. J., & Pachter, L. (2010). Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nature Biotechnology*, *28*(5). https://doi.org/10.1038/nbt.1621

Uezu, A., Kanak, D. J., Bradshaw, T. W. A., Soderblom, E. J., Catavero, C. M., Burette, A. C., Weinberg, R. J., & Soderling, S. H. (2016). Identification of an elaborate complex mediating postsynaptic inhibition. *Science*, *353*(6304). https://doi.org/10.1126/science.aag0821

Ventura, R., & Harris, K. M. (1999). Three-dimensional relationships between hippocampal synapses and astrocytes. *Journal of Neuroscience*, *19*(16). https://doi.org/10.1523/jneurosci.19-16-06897.1999

Victor Nadler, J. (2010). Plasticity of glutamate synaptic mechanisms. *Epilepsia*, *51*(SUPPL. 5). https://doi.org/10.1111/j.1528-1167.2010.02803.x

Virchow, R. (1856). Gesammelte Abhandlungen zur Wissenschaftlichen Medizen. *Frankfurt Am Main: Meidinger U Comp*.

Vogel, C., & Marcotte, E. M. (2012). Insights into the regulation of protein abundance from proteomic and transcriptomic analyses. *Nature Reviews Genetics*, *13*(4). https://doi.org/10.1038/nrg3185

Walz, W. (2000). Role of astrocytes in the clearance of excess extracellular potassium. *Neurochemistry International*, *36*(4–5). https://doi.org/10.1016/S0197-0186(99)00137-0

Wan, Y., Ade, K. K., Caffall, Z., Ilcim Ozlu, M., Eroglu, C., Feng, G., & Calakos, N. (2014). Circuit-selective striatal synaptic dysfunction in the sapap3 knockout mouse model of obsessive-compulsive disorder. *Biological Psychiatry*, *75*(8). https://doi.org/10.1016/j.biopsych.2013.01.008

Welch, J. M., Lu, J., Rodriguiz, R. M., Trotta, N. C., Peca, J., Ding, J. D., Feliciano, C., Chen, M., Adams, J. P., Luo, J., Dudek, S. M., Weinberg, R. J., Calakos, N., Wetsel, W. C., & Feng, G. (2007). Cortico-striatal synaptic defects and OCD-like behaviours in Sapap3-mutant mice. *Nature*, *448*(7156). https://doi.org/10.1038/nature06104

Welch, J. M., Wang, D., & Feng, G. (2004). Differential mRNA expression and protein localization of the SAP90/PSD-95-associated proteins (SAPAPs) in the nervous system of the mouse. *The Journal of Comparative Neurology*, *472*(1), 24–39. https://doi.org/10.1002/cne.20060

Wethmar, K., Smink, J. J., & Leutz, A. (2010). Upstream open reading frames: Molecular switches in (patho)physiology. In *BioEssays* (Vol. 32, Issue 10). https://doi.org/10.1002/bies.201000037

Yang, J. W., Rodrigo, R., Felipo, V., & Lubec, G. (2005). Proteome analysis of primary neurons and astrocytes from rat cerebellum. *Journal of Proteome Research*, *4*(3). https://doi.org/10.1021/pr049774v

Yin, H. H., Mulcare, S. P., Hilário, M. R. F., Clouse, E., Holloway, T., Davis, M. I., Hansson, A. C., Lovinger, D. M., & Costa, R. M. (2009). Dynamic reorganization of striatal circuits during the acquisition and consolidation of a skill. *Nature Neuroscience*, *12*(3). https://doi.org/10.1038/nn.2261

Yu, W., Clyne, M., Khoury, M. J., & Gwinn, M. (2009). Phenopedia and genopedia: Disease-centered and gene-centered views of the evolving knowledge of human genetic associations. In *Bioinformatics* (Vol. 26, Issue 1). https://doi.org/10.1093/bioinformatics/btp618

Yu, X., Nagai, J., Marti-Solano, M., Soto, J. S., Coppola, G., Babu, M. M., & Khakh, B. S. (2020). Context-Specific Striatal Astrocyte Molecular Responses Are Phenotypically Exploitable. *Neuron*, *108*(6). https://doi.org/10.1016/j.neuron.2020.09.021

Yu, X., Taylor, A. M. W., Nagai, J., Golshani, P., Evans, C. J., Coppola, G., & Khakh, B. S. (2018). Reducing Astrocyte Calcium Signaling In Vivo Alters Striatal Microcircuits and Causes Repetitive Behavior. *Neuron*, *99*(6). https://doi.org/10.1016/j.neuron.2018.08.015

Zhang, Y., Chen, K., Sloan, S. A., Bennett, M. L., Scholze, A. R., O'Keeffe, S., Phatnani, H. P., Guarnieri, P., Caneda, C., Ruderisch, N., Deng, S., Liddelow, S. A., Zhang, C., Daneman, R., Maniatis, T., Barres, B. A., & Wu, J. Q. (2014). An RNA-sequencing transcriptome and splicing database of glia, neurons, and vascular cells of the cerebral cortex. *Journal of Neuroscience*, *34*(36). https://doi.org/10.1523/JNEUROSCI.1860-14.2014

Zlatkine, P., Mehul, B., & Magee, A. I. (1997). Retargeting of cytosolic proteins to the plasma membrane by the Lck protein tyrosine kinase dual acylation motif. *Journal of Cell Science*, *110*(5). https://doi.org/10.1242/jcs.110.5.673

Züchner, S., Wendland, J. R., Ashley-Koch, A. E., Collins, A. L., Tran-Viet, K. N., Quinn, K., Timpano, K. C., Cuccaro, M. L., Pericak-Vance, M. A., Steffens, D. C., Krishnan, K. R., Feng, G., & Murphy, D. L. (2009). Multiple rare SAPAP3 missense variants in trichotillomania and OCD. In *Molecular Psychiatry* (Vol. 14, Issue 1). https://doi.org/10.1038/mp.2008.83