Enzyme mediated labeling of synaptic proteins for proteomics and imaging

by

Huaijin Ken Leon Loh

B.S. Chemistry (2009)
Harvey Mudd College

Submitted to the Department of Chemistry in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy

at the

Massachusetts Institute of Technology

June 2016

© 2016 Massachusetts Institute of Technology
All rights reserved

Signature of the Author: ____________________________
Department of Chemistry
April 11, 2016

Certified by: ____________________________
Alice Y. Ting
Ellen Swallow Richards Professor of Chemistry
Thesis supervisor

Accepted by: ____________________________
Robert W. Field
Chairman, Departmental Committee on Graduate Students
This doctoral thesis has been examined by a committee of the Department of Chemistry as follows:

Signature redacted

Joanne Stubbe
Novartis Professor of Chemistry, Professor of Biology

Signature redacted

Alice Y. Ting
Ellen Swallow Richards Professor of Chemistry
Thesis supervisor

Signature redacted

Kay M. Tye
Assistant Professor of Neuroscience
Enzyme mediated labeling of synaptic proteins for proteomics and imaging

by

Huajin Ken Leon Loh

Submitted to the Department of Chemistry on May 30, 2016 in partial fulfillment of the Requirements for the Degree of Doctor of Philosophy

ABSTRACT

The neuronal synapse is one of the most important cellular structures in neuroscience. It is the junction through which signaling occurs between neurons that in aggregate, mediates complex physiological brain function. At a molecular level, this is the site where neurotransmitters are released and recognized, and where remodeling of its protein components occurs to tune the connection strength between neurons, processes crucial to brain functions such as learning and memory.

Despite its importance to brain function, its small size makes synapses difficult to identify by conventional microscopy. Particularly, there remains an incomplete knowledge of the molecular components of the synaptic cleft, the gap space between the pre and postsynaptic neuron, hindering a complete mechanistic understanding of this structure. This is because the cleft is extracellular and non-membrane bound, and hence intractable to biochemical fractionation. Herein, we describe chemical genetic methodologies that allow for the enzymatic labeling of synaptic proteins either for the discovery of new synaptic components, or to elucidate specific trans-synaptic interactions at this inter-cellular junction.

The first part of this thesis describes the application of peroxidase mediated, proximity-dependent biotinylation to map the proteomes of the excitatory and inhibitory synaptic clefts, which mediate distinctively different intercellular signaling between neurons; the former propagates neuronal signals while the latter inhibits them. This work identified numerous novel synaptic proteins and allowed for a comprehensive molecular differentiation of the two synaptic cleft types. Using the proteomes, we identified a specificity factor that mediates accurate matching of inhibitory presynaptic terminals to their post-synaptic cognate receptors. The second half of this thesis describe enzymatic probe ligation methodologies for synaptic proteins, 1) to identify the trans-synaptic interaction of the neurexin-neuroligin complex across the synaptic cleft, and 2) an “eraser” for LplA based probe ligation.

Thesis Supervisor: Alice Y. Ting Title: Ellen Swallow Richards Associate Professor of Chemistry
Acknowledgements:

I would like to thank Alice, my advisor, for so many things, most of all for her enthusiasm and fearless vision for identifying and solving challenging problems. This work would not have been possible without the high expectations she has of the science every Ting lab member can achieve and I can’t think of how many countless times I thought an experiment was impossible, only to reluctantly try it on Alice’s enthusiasm and find it to work out perfectly. I will always be thankful for her mentorship and amazing support. I would also like to thank the members of my thesis committee, Joanne for her ever steadfast and rigorous dedication to science; our conversations at the annual meetings always inspired me and her passion for science is something I strive to always have, and Kay for her kind support and advice, and for being such a great role model and proving that it’s completely possible to balance it all.

I would like to thank the Ting lab for being such an amazing group of scientists and friends. Dan, Tao, Peng and Katie, Wenjing, Hyung woo, Faysal and Sujet for being such awesome lab mentors and teaching me the ropes, Vicky, Kurt, Steph, Jen and Justin for being the best bay to be in, Philipp, Kayvon and Austin for being wonderful to work together with. Jake, Carolyn, Chai, Tess, Yisu, Oom, Monica, Dan Dai, Xin, Anupong, Sam, Phil Zegelbone, Cathy, Jennifer, Mateo, Ozan and Shuo for being such great lab mates and terrific company. Jeff, Marco, Hongik and Yifeng for being such awesome peers who gave me plenty of advice on the science and life in general.

I would like to thank my parents and family for being a never wavering fountain of support, they have always given me the freedom to pursue my interests and I can’t ever be more grateful for the opportunities you have given me.

I’m also really glad to have had Donny, Josh, Andrew, Steve and the Knockout Kings crew who have been such great friends, practice with you guys made it possible to keep coming back to the science with renewed vigor.

Most of all, I would like to thank my wife Sarah, for being the most amazing person in the world. I cannot think of a person who understands me better scientifically and is my soul mate in all aspects of life. None of this work would have been possible without your endless love and support.
Table of contents

Chapter 1: Introduction to proteomics of neuronal synapses --------------- 7
  Introduction--------------------------------------------------------------- 8
  Functional subtypes of synapses------------------------------------------- 9
  Mass spectrometry for subcellular proteomics----------------------------- 12
  Sub-cellular isolation of synapses (synaptosomes)------------------------ 13
  Synaptosome based proteomics--------------------------------------------- 14
  Isolating synaptic proteins by immunopurification------------------------ 17
  Limitations of traditional approaches to purifying synapses------------- 18
  Conclusions---------------------------------------------------------------- 20
  References---------------------------------------------------------------- 21

Chapter 2: Mapping the synaptic cleft by peroxidase mediated biotinylation ----------------------------------------------- 26
  Introduction---------------------------------------------------------------- 27
  Methodologies for promiscuous labeling------------------------------------ 28
  Applying peroxidase mediated biotinylation to the synaptic cleft--------- 32
  Transgene delivery of HRP to neurons-------------------------------------- 35
  BxxP - cell impermeable probe for biotinylation------------------------- 37
  HRP mediated biotinylation of synapses------------------------------------ 40
  Ultrastructure localization of HRP fusions in the synapse by electron microscopy -49
  Streptavidin bead enrichment of synapses for proteomics----------------- 52
  Preliminary synaptic cleftome using Neuroligin 1------------------------ 55
  Dissociating the PSD for pulldowns-------------------------------------- 68
  Targeting the excitatory and inhibitory synaptic clefts----------------- 71
  Conclusions---------------------------------------------------------------- 80
  Materials and Methods---------------------------------------------------- 81
  References---------------------------------------------------------------- 93

Chapter 3: The excitatory and inhibitory synaptic cleft proteomes------- 99
  Introduction---------------------------------------------------------------- 100
  Proteomic mapping of the excitatory and inhibitory synaptic cleft------- 100
  Determining cutoffs the final synaptic cleft proteomes------------------ 106
  Characterization of the excitatory and inhibitory synaptic cleft proteomes ----- 113
  Depth of coverage analysis of synaptic cleftomes------------------------ 117
  Validation of proteomic data--------------------------------------------- 120
  Biological Implications of new assignments------------------------------ 125
  Conclusions---------------------------------------------------------------- 129
  Materials and Methods---------------------------------------------------- 131
  References---------------------------------------------------------------- 137

Chapter 4: Identification of MDGA2 as an inhibitory synapse specificity factor ----------------------------------------------------- 145
  Introduction---------------------------------------------------------------- 146
  MDGAs are localized at synapses------------------------------------------- 148
  Synaptogenesis assay to identify and compare MDGA function------------- 150
  Loss of function experiments to determine MDGAs effect on synapse density --- 154
  MDGAs domains responsible for its synaptic localization---------------- 156
  Model for MDGA function and implications------------------------------- 159
  Conclusions---------------------------------------------------------------- 161
<table>
<thead>
<tr>
<th>Chapter 5: Mapping the Axon initial segment using APEX</th>
<th>168</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>169</td>
</tr>
<tr>
<td>Targeting APEX to the Axon initial segment</td>
<td>170</td>
</tr>
<tr>
<td>Mapping the proteome of the Axon initial segment</td>
<td>176</td>
</tr>
<tr>
<td>Optimizations for subsequent proteomic experiments</td>
<td>183</td>
</tr>
<tr>
<td>Conclusions</td>
<td>185</td>
</tr>
<tr>
<td>Materials and Methods</td>
<td>186</td>
</tr>
<tr>
<td>References</td>
<td>188</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Chapter 6: Enzymatic probe ligation of neurexin-neuroligin interactions in neurons</th>
<th>191</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>192</td>
</tr>
<tr>
<td>Methodologies for labeling trans-synaptic interactions</td>
<td>192</td>
</tr>
<tr>
<td>Identifying optimal sites for fusing biotin-ligase in neurexin</td>
<td>200</td>
</tr>
<tr>
<td>Biotin-ligase dependent labeling of the neurexin-neuroligin complex in neurons</td>
<td>203</td>
</tr>
<tr>
<td>Lipoic acid ligase (LpIA) dependent labeling of the neurexin-neuroligin complex</td>
<td>208</td>
</tr>
<tr>
<td>Conclusions</td>
<td>217</td>
</tr>
<tr>
<td>Materials and Methods</td>
<td>220</td>
</tr>
<tr>
<td>References</td>
<td>225</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Chapter 7: Engineering lipoamidase as an eraser for LpIA dependent labeling</th>
<th>230</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>231</td>
</tr>
<tr>
<td>Identifying an enzyme to reverse LpIA dependent labeling</td>
<td>232</td>
</tr>
<tr>
<td>Purification of Lipoamidase</td>
<td>234</td>
</tr>
<tr>
<td>Engineering a coumarin amidase from Lpa</td>
<td>247</td>
</tr>
<tr>
<td>Lipoamidase has resorufin amidase activity</td>
<td>254</td>
</tr>
<tr>
<td>Conclusions</td>
<td>257</td>
</tr>
<tr>
<td>Materials and Methods</td>
<td>258</td>
</tr>
<tr>
<td>References</td>
<td>261</td>
</tr>
</tbody>
</table>
Chapter 1: Introduction to proteomics of neuronal synapses
Introduction

We owe our ability to think, to learn, to remember, to perceive our surroundings and so many other cognitive functions to the complex organ that is the mammalian brain. Much of this is accomplished by interconnected nerve cells within the brain where each neuron acting as a basic unit that processes and passes on information between neurons that have been assembled into an anatomical circuit. Central to this information transfer, which can encode for a complex behavior or memory, are highly specialized subcellular structures called synapses that are formed at sites of contact between neurons, that are one of the most important sub-cellular structures in neuroscience. In the vertebrate central nervous system (CNS), neuronal synapses are primarily chemical synapses (in comparison to the gap-junction based electrical synapse (Purves et al., 2001)), which is comprised of 3 distinct subcellular regions (shown in Fig. 1-1): 1) the presynaptic terminal that comes from the axon of the signal sending neuron, this contains neurotransmitter containing synaptic vesicles that upon stimulation by an electrical signal (caused by calcium influx into the terminal), can fuse with the plasma membrane and unload their neurotransmitter cargo into the ~30nm space between the two neurons called the 2) synaptic cleft. Receptors that bind these neurotransmitters can be found on the membrane surface of the dendrites of the neuron receiving the signal, in a region called 3) the post-synaptic density (PSD) that is densely packed with receptors, scaffold proteins and numerous others involved in signaling. Additionally, the presynaptic terminal can be further subdivided into the membrane region just juxtaposed to the synaptic cleft; this region is called the active zone and contains specialized machinery to enable the rapid fusion and exocytosis of neurotransmitter containing vesicles with the plasma membrane during an action potential. These subcellular compartments serve to convert electrical signaling (in the form of an action potential in the presynaptic neuron) into a chemical signal (neurotransmitters that bind to receptors) back into an electrical signal (by the action of cation permeability into the cell by ionotropic receptors).
Functional subtypes of synapses

Chemical synapses can be further subdivided into different classes based on the neurotransmitter utilized. A neurotransmitter must meet the following four criteria to be considered as such (Kandel et al. 2000): 1) it is synthesized by the presynaptic neuron, 2) it is present in the presynaptic terminal, and released in sufficient amounts to have an effect on the post-synaptic cell, 3) when added exogenously in physiologically relevant concentrations, it recapitulates the effect of the endogenously released transmitter and lastly, 4) there exists a mechanism for its removal from the synaptic cleft, the site of its action. For this thesis, we will primarily focus on synapses using two types of neurotransmitters; glutamate and GABA (gamma amino-butyric acid) that are the most common neurotransmitters used in synapses derived from neurons in the cortex and are abundantly found in cultured cortical neurons. Functionally, the synapse types that use each of these
neurotransmitters have highly divergent but complementary functions that are essential towards normal brain function.

Glutamatergic synapses use glutamate and are the most common synapses found in the CNS. Presynaptic terminals concentrate glutamate into their synaptic vesicles by using one of three vesicular glutamate transporters (vGLUT1-3), which also serve as markers of these synapses. Glutamate when released into the synaptic cleft, binds to glutamate receptors (Traynelis et al., 2010) such as AMPA and NMDA receptors on the surface of the signal receiving post-synaptic neuron, which in turn are cation-selective ion channels that allow calcium and sodium ions into the neuron upon their binding. This mode of action serves to act towards depolarizing the post-synaptic neuron, increasing the propensity of an action potential occurring, and propagating the signal between neurons. In this manner, glutamatergic synapses can be considered to be “excitatory” and will be referred interchangeably as such.

GABAergic synapses use GABA (gamma amino-butyric acid) as their neurotransmitter and are important for proper brain function because they perform the opposing role from excitatory synapses; inhibiting an action potential from occurring. GABA is packaged into synaptic vesicles by the action of the inhibitory amino acid transporter vGAT, which is also localized exclusively to inhibitory presynapses. Inhibition of the post-synaptic neurons occurs when GABA is released from the presynaptic neuron into the synaptic cleft, binds to ionotrophic GABA receptors such as the GABAA receptor family (Chebib and Johnston, 1999; Sigel and Steinmann, 2012), which allow chloride ions to pass through the membrane and into the cell when GABA is bound. This influx of chloride serves to hyperpolarize the post-synaptic neuron, making it more difficult for an action potential to occur and inhibiting the signal from being propagated. In this way one can view GABAergic synapses as “inhibitory”.

Morphologically, these synapses appear differently when observed at the ultrastructural level. Gray identified two types of synapses in the cortex by electron microscopy, the Type 1 or asymmetric synapses (found to be the above described glutamatergic and excitatory) and the Type 2 or symmetric synapses
(found to be GABAergic or inhibitory) (Gray EG (GRAY, 1959)1959) (Fig. 1-2). The former were found primarily on neuronal dendritic shafts and spines while the latter were found on dendritic shafts and neuronal cell bodies. The Type 1 presynapses often extended from axons that had round or extended vesicles, apposed across the synaptic cleft by a thickening of electron dense material (PSD) giving the impression of the structure as asymmetric. Type 2 on the other hand, had pre and post-synaptic membranes that appeared more or less parallel, often contained vesicles that were elongated in shape and noticeably did not contain a PSD (Colonnier 1968).

Figure 1-2. **Type 1 and Type II synapses as described by Gray.** Figure is adapted from (Harris and Weinberg, 2012). Asymmetric synapses can be observed in D (green arrows) along the shaft of a dendrite containing a mitochondrion (mito). These have comparable thicknesses to the perforated PSDs on dendritic spines in the same field (marked by red arrow). E and F show adjacent sections through a symmetric synapse (orange arrows) on a dendritic shaft. Note the difference in staining as a reflection of protein density.

Most importantly, these two synapse types have highly divergent functions and a balance of excitation and inhibition is absolutely necessary for proper brain function. Many of the protein components of each of these synapse types themselves are related to neurologic and cognitive dysfunction and some examples of proteins found within these synapses that are linked to psychiatric diseases such as autism spectrum disorders include the synaptic adhesion
neuroligins and neurexins (Südhof, 2008), while mGluR5 is associated with Fragile X syndrome.

**Mass spectrometry for subcellular proteomics**

Since the molecular content of these structures is fundamental towards proper synapse function, which in turn translates to proper brain function, it is really important to identify what molecular players that reside within each of these substructures. To identify these proteins one can take on a gene by gene targeted discovery or use a large scale, unbiased approach with profiling methodologies such as mass-spectrometry based proteomics (Bayès and Grant, 2009) towards solving this problem. The term proteome was derived as an analogous term to genome, and represents a collection of all the proteins that could be translated from the genome. Proteomics has been incredibly powerful at cataloguing the molecular components in a region of interest that you want to study, allowing one to make informed hypothesis of the proteins that might be relevant to the function of that region of interest. A simplified traditional workflow for this is shown in Fig. 1-3: first a region of interest is isolated; this could be a cell type, an organelle or even a protein complex, often by using traditional biochemical fractionation. Once proteins are isolated, proteins mixtures are fractionated further, often by gel electrophoresis, and protein bands are cut out, and proteins eluted and proteolytically digested. These proteins are then identified using liquid chromatography tandem mass spectrometry (LC/MS/MS) that acts as a high-throughput protein sequencer. The key step, and also the most difficult for obtaining an accurate and reliable dataset is the purification step because isolating a region of interest is not always feasible, while contaminants introduced during purification generate false positives that are difficult to distinguish from true positives.
Sub-cellular isolation of synapses (synaptosomes)

Synapses were first isolated biochemically from brain tissue by Whittaker in 1958 (Hebb and Whittaker, 1958) (This is a review and perspective of this work (Whittaker, 1993)). These were nerve terminals whose membranes had resealed upon separation from their axons and dendrites by liquid shear, when brain tissue is homogenized under non-ionic and iso-osmotic conditions in a sucrose suspension media. A series of differential centrifugation and sucrose-density gradient centrifugations were then carried out to obtain what Whittaker termed “synaptosomes”. Variations of this purification protocol exist; such as using a hyperosmotic gradients or iso-osmotic Ficoll gradients but in general the purest synaptosomes are derived from the original protocol (Whittaker 1993). When viewed under an electron microscope, these structures contained many of the structural features observed in pre and post-synapses from brain tissue; free synaptic vesicles (both small clear ones and the larger dense core vesicles), usually one or two mitochondria, an active zone and the electron dense “darkening” that constitutes part of the post-synaptic density. An electron microscope image of these synaptosomes can be viewed here (Fig. 1-4). These synaptosomes have been instrumental towards the understanding of synaptic...
structure, and also function; for example synaptosomes enabled the understanding of amino acid uptake into synaptic vesicles, and because they contain most of the components of pre-synapses, they are competent for depolarization and release neurotransmitter in a calcium dependent manner. Most importantly, because this provides a scalable method towards obtaining synapses biochemically, a route to which sufficiently pure material for mass spectrometry proteomics is possible, to elucidate the components of the synapse.

![Figure 1-4. Whittaker’s synaptosomes as viewed by Gray using electron microscopy.](image)

Figure is adapted from (GRAY and WHITTAKER, 1962). Synaptic vesicles are visible within the synaptosome often accompanied by a mitochondria. A protein rich darkened membrane corresponding to the post-synaptic membrane (psm in image) is often found attached to the synaptosome.

**Synaptosome based proteomics**

Since its biochemical isolation, numerous synaptosome based proteomic experiments have been carried out to identify the components of the neuronal synapse. These are reviewed extensively here (Abul-Husn and Devi, 2006; Bai and Witzmann, 2007; Bayés and Grant, 2009). The first proteomic analyses (Boyd-Kimball et al., 2005; Schrmpf et al., 2005; Witzmann et al., 2005) of the synaptosome identified greater than 1000 unique proteins (Grant, 2006), providing the first molecular catalog of the synapse. Synaptosome proteomics has also been used to characterize protein expression, on synaptosomes derived from a pathological phenotype, such as following a nerve injury (Singh et al., 2009) and has also been applied towards understanding changes in synaptic
content following drugs that target neurological diseases such as schizophrenia (Ji et al., 2009). Protein subsets describing the sub-compartments of the synapse have been extracted using modifications to the synaptosome fractionation protocol, by leveraging biochemical properties of these sub-compartments.

The PSD was first isolated from synaptosomes (Cohen et al., 1977; Cotman et al., 1974; Davis and Bloom, 1973; Matus and Taff-Jones, 1978) in the 1970s by groups who discovered that the PSD could be separated from synaptosomes because it was detergent resistant. Application of detergents such as triton X-100 (Carlin et al., 1980) and deoxycholate would leave the PSD relatively intact but would separate synaptic membranes (including the presynaptic terminal and synaptic vesicles) from the PSD. A centrifugation step following extraction would then allow for purification of the membrane free PSD.

The application of mass spectrometry proteomics allowed for subsequent identification of more than 2000 proteins from seven different studies (reviewed in Bayés and Grant, 2009; Sheng and Hoogenraad, 2007; Walikonis et al., 2000) although only less than half (466) of the proteins identified were identified in all studies, suggesting that some of these proteins could be false positives or context specific, and comprise a "core" PSD machinery common to all excitatory synapses (Collins et al., 2006). Proteins identified in the PSD included cytoskeletal proteins, numerous glutamate receptors such as the N-methyl-D-aspartate (NMDA) receptor and AMPA receptor complex and their associated proteins, calmodulin dependent protein kinase II (CaMKII) signaling components, and scaffolding components such as members of the membrane-associated guanylate kinase homologues (MAGUKs). Many of these components were known to reside within the glutamatergic synapse, suggesting that these PSD purifications were mainly of the excitatory nature. Many of these proteins however, remain uncharacterized and their function within the synapse remains unknown. Such lists are also likely to contain false positives due to the impurity associated with subcellular fractionation (e.g., contamination by mitochondria and other organelles) (Sheng and Kim, 2011), so it is likely that the number of unique proteins at PSDs is closer to 500 than 2000. A brief inspection of proteomic
preparations of PSDs reveals that at least 10-20% of proteins are also of mitochondrial and nuclear origin (Bayés et al., 2012; Collins et al., 2006).

Synaptic vesicle enriched fractions have been obtained by modifications of the classic Whittaker protocol (Nagy et al., 1976) and there are protocols that can generate morphologically homogenous synaptic vesicles with approximately 95% of vesicles containing major synaptic vesicle proteins such as synaptophysin (Huttner et al., 1983). This involves the generation of synaptosomes using the Whittaker protocol, followed by an osmotic lysis and centrifugation steps to remove synaptic membranes components, ending with a chromatography step on controlled pore glass beads to isolate synaptic vesicles. This has been used to build a highly quantitative model of the average synaptic vesicle (Takamori et al., 2006) by mass spectrometry. One feature of this proteomics experiment was the identification of vesicular neurotransmitter transporters such as vGlut1-2, vGAT and vAChT and vMat2 revealing that the synaptic vesicles isolated were derived from heterogeneous neurotransmitter subtypes. Another issue associated with this protocol is the contamination of proteins from organelles other than synaptic vesicles (such as mitochondria), this has been improved upon by the introduction of an immunoisolation step using antibodies to known synaptic vesicle proteins (Morciano et al., 2005).

Of the three cytoplasmic sub-compartments, the synaptic active zone has probably been the most difficult to isolate cleanly, because despite detergent extraction removing many of the membrane bound components that form part of the exocytosis machinery on the presynaptic side, within the detergent insoluble fraction that remains and is thought to comprise the PSD, many presynaptic membrane proteins remain because of the tight connection between the pre and post-synaptic sides (Langnaese et al., 1996). Various protocols have been used to try to separate presynaptic-membranes from post-synaptic ones, such as differential pH extractions or treatment with urea although most of these fail to detect most known active zone components. To circumvent this issue, Jahn and co-workers employed an additional step of mild proteolysis of synaptosomes to dissociate the pre and post-synaptic membranes assuming that protease
digestion would only work on proteins not protected within a membrane (such as within the synaptic cleft), followed by the standard osmotic lysis and centrifugation used to isolate synaptic vesicles. The additional immunoisolation step used by Morciano and co-workers above is then used to isolate both synaptic vesicles and docked complexes comprising the active zone. This protocol allowed them to isolate and identify almost all known active zone proteins, and numerous ion channels (such as the voltage gated calcium channel subunits) and transporters expected to belong to the presynaptic membrane (Boyken et al., 2013).

Isolating synaptic proteins by immunopurification

Synaptic proteins can also be isolated by immunoprecipitation or affinity purification. This is the application of highly specific antibodies towards the isolation of specific proteins from a complex protein mixture. This strategy is particularly suitable for isolating the PSD because the excitatory post-synapse is a complex that is highly cross-linked (Carlin et al., 1980) and detergent resistant, and is likely to retain its complex components even under more purification and washing conditions. Grant and co-workers isolated mouse protein PSD complexes via the scaffold protein PSD95, a member of the MAGUKs and highly abundant member of the excitatory PSD. PSD95 was genetically modified with two epitope sequences for tandem affinity purification (TAP tags) (Fernández et al., 2009), and the isolated complex contained 301 proteins, of which 118 were known components of the postsynaptic excitatory synapse, such as glutamate receptors and potassium channels. Other PSD complexes purified by immunoprecipitation include the protein complexes associated with PSD-95 and NMDA receptors each of which has led to the identification of 288 and 77 proteins, respectively (Dosemeci et al., 2007; Husi et al., 2000), and many of these proteins themselves are overlapping members of the core set of proteins in the PSD described earlier.
Immunoprecipitation focused strategies have also been applied towards the inhibitory post-synapses that have very different biochemical properties from the excitatory post-synaptic density. The symmetric type II synapse (inhibitory synapse) is named as such (GRAY, 1959) because it does not contain a protein dense post-synaptic density; as such identifying its components has been difficult since there are no biochemical sources available for proteomic identification (typical PSD purification protocols cannot be applied to it since it is not detergent resistant), and identifying inhibitory synapse proteins requires careful and painstaking, one by one identification and validation. The Heintz and Craig labs each employed an immunoprecipitation strategy using different inhibitory synapse baits. The Heintz lab attached a eGFP tag (that could serve as an immunopurification handle) to the N-terminus of the GABA$_A$ R $\alpha$1 subunit, since GABA$_A$ R $\alpha$1 is inserted into the majority of GABA$_A$ receptors, the main receptor in GABAergic inhibitory synapses (Heller et al., 2012). The Craig lab fused a His6-FLAG-YFP tag onto Neuroligin 2 as their immuno-purification handle, an inhibitory post-synapse specific protein, that functions in cell adhesion and synaptic organization of synapses, via their binding to the neurexin family of proteins on the presynapse (Kang et al., 2014). While these interactomes can be quite specific (Heller et al., 2012), the majority of proteins resident in the inhibitory synapse are missed because they do not interact directly and/or stably with the bait proteins, limiting their identification.

Limitations of traditional approaches to purifying synapses

One major limitation of any proteomic experiment based on the purification of synaptosomes (which includes almost all protocols described earlier) is that there is no evidence to suggest that synaptosomes are specific for any subtype of synapse or transmitter, and the only way to obtain a homogenous mixture of synaptosomes that use the same neurotransmitter is to purify synaptosomes from tissue that are relatively homogenous in their neurotransmitter content, such as the electric organ of the electric eel, which are relatively homogenous in
cholinergic synapses (Dowdall and Whittaker, 1973). This however remains a daunting task with regards to synapses in the CNS that are highly heterogeneous in their subtypes (neurotransmitter wise) and almost all neurons are known to receive both excitatory and inhibitory input. A study by the Herzog lab sorted for vGlut1-GFP positive synaptosomes using fluorescence activated cell sorting (FACs) to enrich for proteins enriched in glutamatergic synapses. Although they were able to enrich for glutamatergic components and deplete extrasynaptic and glial contaminants (Biesemann et al., 2014), GABAergic synapse components were still found within their proteomic dataset, highlighting the challenges with effectively sorting between synaptosomes of different neurotransmitter content.

Moreover, typical synaptosome preparations are not actually very pure; a quantitative EM study of synaptosomes prepared by variations of the classical sucrose gradient protocol revealed that only ~50% of all structures observed were actually synaptosomes, and that numerous contaminants such as myelin and mitochondria were often copurified together with the structures (~10%) (Dodd et al., 1981). The rest of the structures were thought to be of neuronal and glial origin (Henn et al., 1976). Even the synaptosomes that had been sorted by FACs contained ~ 20% proteins that are of non-neuronal glial origin (Biesemann et al., 2014). As a result, most proteomes of synaptosomes are likely to contain a large number of false positives, a challenge that limits the utility of these proteomes due to the laborious follow up validation necessary for proving these proteins are synaptic.

As mentioned earlier, inhibitory synapses do not have a PSD; hence it is currently impossible to isolate the post-synaptic fraction of GABAergic terminals. It must be noted however, that GABA_{A} receptors and gephyrin are sometimes isolated together in PSD or synaptosome purifications (Bayés et al., 2012; Collins et al., 2006), but most of the time these are incomplete isolations of known GABAergic post-synapses (many expected subunits missing), and certain known GABAergic adhesion proteins such as the inhibitory synapse adhesion protein Slitrk3 has never been isolated within synaptosome or PSD preparations.
None of the approaches described above are a direct purification of the synaptic cleft; because it is extracellular and non-membrane bound. Components that are not tethered to a membrane, such as secreted proteins are inevitably lost as a result. The mechanism by which cells are lysed disrupts compartments and complexes resulting in a loss of spatial information. Furthermore, contaminants such as mitochondrial proteins (Sheng and Kim, 2011) and the most abundant components of the cell such as protein chaperones, histones and ribosomes are always found within subcellular fractionated proteomic preparations (Mellacheruvu et al., 2013). Because the purifications themselves take from hours to days, it is also very plausible that protein components can be degraded and post-translational modifications such as phosphorylation can be lost as well. Most frustratingly, many compartments & complexes such as the synaptic cleft simply cannot be purified. Because so many synaptic cleft proteins correspond to human disease genes and drug targets for therapeutics that modulate cognitive function, a catalog of these proteins would actually be very beneficial for identifying the molecular players involved in synaptic transmission and specificity.

Conclusions

Here we have reviewed the difficulties associated with biochemical purification of synapses towards identifying proteins residing within synapses. In this thesis, we will discuss the development of enzymatic protein labeling technologies (such as a peroxidase mediated biotinylation) that allows us to label proteins residing within the excitatory and inhibitory synaptic clefts (Chapter 2). This has been applied successfully towards mapping the proteomes of these two previously unpurifiable subcellular regions (Chapter 3). We have used these proteomes to discover novel synaptic proteins, while also revealing localizations of known synaptic proteins thought to be exclusive to one synaptic subtype, within another functionally divergent synaptic cleft type. In Chapter 4, we use our proteomic datasets to identify and follow up on the function of a protein family (MDGAs) residing in the inhibitory synaptic cleft. Chapter 5 discusses the
application of peroxidase mediated biotinylation towards labeling proteins in the axon initial segment, another subcellular and non-purifiable region that is the signal integration hub of the neuron, while Chapter 6 and 7 will discuss the development of two enzyme mediated methodologies for reversible specific labeling and removal of functional handles onto synaptic proteins.

References


Chapter 2: Mapping the synaptic cleft by peroxidase mediated biotinylation

Sections of this chapter are excerpts taken from a submitted manuscript: K. H. Loh, P. S. Stawski, N. D. Udeshi, T. Svinkina, T. Deerinck, M. H. Ellisman, S. A. Carr, and A. Y. Ting. Proteomic maps of the excitatory and inhibitory synaptic clefts reveal specificity factor for pre-to-post synaptic matching. Electron microscopy was kindly performed in collaboration with Philipp Stawski and Tom Deerinck in Mark Ellisman’s lab at UCSD. Proteomics was performed in collaboration with Namrata Udeshi and Tanya Svinkina in Steven Carr’s lab at the Broad Institute. Victoria Hung and Kurt Cox provided invaluable advice on analysis and developing the methodology.
Introduction

As discussed in Chapter 1, it has been impossible to purify the substructure that comprises the synaptic cleft, because it is incompatible with standard synaptosome protocols and their derivatives. One could however, approximate the content within the synaptic cleft by looking at surface exposed candidate proteins identified in the synaptosome purifications, however this still remains an indirect approach for identifying cleft proteins, and does not distinguish the neurotransmitter type of synapse these proteins function within.

We envisioned that a methodology that can help us tag proteins within the synaptic cleft with biotin, would provide us with a biochemical handle which we can use to isolate synaptic cleft proteins, while circumventing the need for classical organelle purification (Fig. 2-1). Streptavidin, a protein purified from Streptomyces avidinii is a protein that binds biotin with a $K_d$ of around $10^{-14}$ M (Green, 1975), and has been employed extensively for affinity purification based strategies for isolating proteins and nucleotides tagged with biotin. This interaction is one of the strongest non-covalent biochemical interactions and we thought we could leverage this to enrich very small amounts of biotinylated protein from a large complex milieu.

---

**Figure 2-1. Overview of scheme for mapping the synaptic cleft without organelle purification.** A promiscuous labeling enzyme is first genetically targeted to the cleft. The grey
shapes are endogenous proteins residing inside and outside the synapse. To initiate labeling, a biotin containing substrate is added to the live neurons for 1 minute to initiate labeling. The enzyme catalyzes the formation of reactive biotin species, which covalently tags proximal endogenous proteins. Subsequently, neurons are lysed and biotinylated proteins are isolated using streptavidin (SA) beads for identification by mass spectrometry (MS).

Methodologies for promiscuous labeling

Key to this approach is the need for a promiscuous labeling enzyme; which can catalyze the formation of a reactive biotin tag, that diffuses from its active site to label proteins proximal to the enzyme. Ideally this chemical species has the following characteristics; it is highly reactive towards functionalities widely available on the surface of proteins and short lived limiting the labeling radius to within nanometers of the enzyme and not much further, so that it can label a structure as small as the synaptic cleft (Schikorski and Stevens, 1997).

Figure 2-2. Promiscuous labeling strategy. An enzyme is used to generate a reactive biotin containing species, that is reactive enough that it only labels proximal endogenous proteins nanometers away from the enzyme and not further.

Promiscuous biotinylation using pBirA is incompatible with synaptic cleft proteomics

One candidate could be the *Escherichia coli* biotin ligase mutant R118G (promiscuous BirA or pBirA), first suggested by the Cronan group as an enzyme that could be used for proximity dependent promiscuous protein biotinylation (Choi-Rhee et al., 2004). pBirA generates a biotin-AMP anhydride, that can diffuse from its active site to tag proteins at lysine residues (Fig. 2-2) with a covalent biotin modification. pBirA is able to function this way because it contains a mutation within a loop that is normally disordered, that becomes ordered when biotin is bound, and this mutation leads to defective binding of both biotin and

28
biotin-AMP once it has been formed (Weaver et al., 2001). Roux et al. have extended the application of pBirA to biotinylation in living mammalian cells (Roux et al., 2012), resulting in biotinylation of endogenous proteins that can be selectively enriched using streptavidin beads for mass spectrometry, and this methodology is known as BioID. This was applied to the nuclear lamina by fusing pBirA to the nuclear lamin A protein. The authors used this technique to identify about 100 candidates including several previously uncharacterized nuclear lamina proteins such as SLAP75.

This enzyme is not suitable for labeling of synaptic cleft proteins because of poor spatial resolution and slow kinetics. The kinetics of biotin-AMP generation are \(0.24 \text{ s}^{-1}\) (Kwon and Beckett, 2000), and requires a minimum of 6 hours of biotin-incubation to generate sufficient labeling for mass spectrometry analysis (Roux et al., 2012). This long labeling time window could give rise to undesirable functional consequences on proteins that have been post-translationally modified with biotin by pBirA and would not be suitable for studying a structure that is dynamically regulated like the synapse. Furthermore the biotin-AMP half-life is estimated to be around 30 minutes in water based on comparisons of the stability of acyl adenlyates such as glutaminyl-adenylate, whose rate constant for hydrolysis is about \((0.7 \times 10^{-3} \text{ s}^{-1})\) (Gruic-Sovulj et al., 2005). This suggests its labeling radius would be large enough to transverse an entire neuron as well based on simple diffusion. If we approximate the synaptic cleft as being a structure that is 50nm across the cleft, while \(~200\)nm wide, such a labeling protocol is not sufficient to keep the labeling radius within the size of such a structure, and non-synaptic protein labeling will occur within the labeling window. Additionally, since the labeling is dependent on the presence of deprotonated lysines (pKa 10.5), this reaction is not favored at physiological pHs.
Figure 2-3. Promiscuous biotin ligase as a candidate enzyme for proteomic tagging. pBirA generates and releases biotin-adenylate from its active site, that can label neighboring proteins on surface exposed lysine residues.

The Ting lab has recently engineered a promiscuous labeling enzyme that fits the criteria necessary for our approach. This enzyme is a heme peroxidase called APEX (Martell et al., 2012), or APEX2 for the most evolved active variant (Lam et al., 2014). APEX is a 27 kDa protein derived from soybean ascorbate peroxidase (APX), which is isolated from plants that acts to prevent cytosolic accumulation of $\text{H}_2\text{O}_2$ by catalyzing its reduction to $\text{H}_2\text{O}$ while oxidizing ascorbate. Dr. Jeffery Martell, a graduate student in Ting lab engineered APX to become a monomer from its native dimer state by introducing 2 mutations (K14D and E112K) at the dimeric interface, and a mutation within the active site (W41F) that improves its activity by about 8-fold for oxidizing a variety of aromatic substrates (Martell et al., 2012). APEX2 is the latest most active variant that a graduate student Dr. Stephanie Lam, used yeast-display (Lam et al., 2014) to evolve, introducing a single mutant (A134P) into APEX, that makes it more resistant to reversible $\text{H}_2\text{O}_2$ dependent inhibition via Compound III (Lam et al., 2014). APEX is also active in all subcellular compartments and at most physiological pHs tested (Martell et al., 2012).

APEX can be used for spatially restricted proteomic tagging

APEX’s ability to catalyze, in the presence of the co-oxidant $\text{H}_2\text{O}_2$, the sequential $1e^-$ oxidation of a promiscuous variety of aromatic substrates such as phenols and anilines[ref], together with other characteristics mentioned above makes it an appropriate choice as our promiscuous labeling enzyme. APEX can be used to generate a highly reactive biotin tag appended to a phenoxyl radical (Fig. 2-4) (biotin-phenol also abbreviated as BP), and these phenoxyl radicals are able to react with tyrosine side chains and other $e^-$ rich side chains such as tryptophans, cysteines and histidines (Rhee et al., 2013) on endogenous proteins that are proximal to APEX. This labeling chemistry is also used in the tyramide signal amplification kit sold by Invitrogen (Bobrow et al., 1989), used to amplify
antibody dependent fluorescence staining. Because this radical is highly reactive, and also relatively short lived (half life <1ms) (Mortensen and Skibsted, 1997), these biotin-phenoxyl radicals have a much shorter labeling radius compared to the biotin-AMP formed by pBirA. Furthermore, the kinetics of generating these radicals is extremely fast (Kcat/Km = 10^6 mol^{-1}s^{-1}, (Lam et al., 2014)), and the labeling time can be reduced to as short as 1 min compared to from 6-24 hours for pBirA (in terms of amount of biotinylated material generated). Once proteins have been tagged with biotin, they can then be enriched using streptavidin coated magnetic beads for mass spectrometry based identification. The Ting lab has used APEX to map the proteomes of the mitochondrial matrix(Rhee et al., 2013), which is a completely membrane-bound organelle, and the inner mitochondrial membrane space(Hung et al., 2014) (IMS), which is membrane bound, but semi-porous with the cytosol and allows proteins of <10kDa to pass through.

APEX is also a genetically encoded tag for electron microscopy

APEX is also able to function as a GFP-equivalent of a localization tag for electron microscopy (EM), which can visualize sub-cellular structures down to 1nm in resolution in fixed cells. This is really suitable for visualizing synapses or proteins at synapses, because the size of synapses is very close to the Abbe’s limit of defraction, 250nm (Abbe, 1873). Contrast for EM is often generated by decorating cellular targets of interest with gold nanoparticles and quantum dots, or most commonly by the formation of electron dense deposits. These deposits can be made by polymerizing diamino-benzidine (DAB) (Connolly et al., 1994; Li et al., 2010; De Mey et al., 1981a) either by introducing to the target peroxidases such as horse-radish peroxidase (HRP) conjugated to antibodies (Adams, 1992; Mesulam, 1978; De Mey et al., 1981b), or photosensitzers that generate singlet oxygen with light irradiation(Shu et al., 2011) . The polymer formed, in turn recruits osmium to give the electron dense deposits that act as contrast. APEX is able to act as an electron microscopy tag because it is able to catalyze the polymerization of DAB, and an additional advantage over classical immuno-
labeling strategies for introducing peroxidases for DAB polymerization or gold nanoparticles is the absence of the use of harsh detergents to permeabilize cells, which is pertubative of ultrastructure and causes sample quality to be poorer (Martell et al., 2012).

**HRP can replace APEX in the secretory pathway**

HRP is another peroxidase that is able to catalyze the same reactions (biotin-phenol oxidation to phenoxy radicals, and polymerization of DAB). HRP is only slightly larger than APEX (34 kDa vs 27 kDa), and can replace APEX in the secretory pathway or on the cell surface such as within the synaptic cleft; its peroxidase activity is restricted to those compartments because it contains 4 intramolecular disulfide bonds, in addition to 2 Ca^{2+} ions, and 9 glycosylation sites (Martell et al., 2012). This results in improper folding of HRP within the reducing environment of the cytosol. HRP is also natively a monomer and may be preferable to APEX in these situations because it is more active and because it is less susceptible to H_{2}O_{2} inactivation. HRP also has a similar \( K_{cat}/K_{m} \) to APEX2 for oxidizing these aromatic substrates (\( K_{cat}/K_{m} = 10^{6} \text{ mol}^{-1}\text{s}^{-1} \)).

---

**Figure 2-4. APEX2 can be used for proteomics or electron microscopy.** APEX2 can catalyze both the generation of a DAB polymer for electron microscopy and biotin-phenoxy radicals for labeling endogenous proteins for proteomics.

---

**Applying peroxidase mediated biotinylation to the synaptic cleft**

Since the synaptic cleft is on the cell surface, we decided to select the more active of the two peroxidases (HRP) as a starting point for our promiscuous labeling strategy. HRP has previously been demonstrated to be active within
neurons for DAB and had been used as an EM tag (Li et al., 2010; Schikorski et al., 2007; Watts et al., 2004) for mapping connectivity. We decided to test if HRP was still able to catalyze biotin-phenol labeling of proteins on the neuronal cell surface and to develop labeling conditions that would work well in anticipation of applying this eventually to the synaptic cleft. HRP was fused to the transmembrane domains of various surface proteins including the transmembrane domain of platelet-derived growth factor receptor PDGFR protein (HRP-TM), and we expressed this in cultured rat cortical neurons (Kaech and Banker, 2006) under the CAG (Jun-ichi et al., 1989) (chicken beta actin, cytomegalovirus (CMV) enhancer) promoter (a strong neuronal expression promoter) by lipofecting these neurons with lipofectamine, a commonly used cationic lipid reagent used for introducing plasmids into mammalian cells (Dalby et al., 2004). Biotin-phenol labeling of endogenous proteins was detected by neutravidin (Marttila et al., 2000) (a neutrally charged streptavidin) fluorophore conjugates and imaged by fluorescence microscopy. Labeling conditions (100uM biotin-phenol probe and 1mM H₂O₂ for 1 min) previously used in HEK cells (Rhee et al., 2013) were found to be sufficient for obtaining good labeling (Fig. 2-5). Furthermore neuronal morphology was found to look normal following labeling (such as preservation of dendritic spine structures) and cell blebbing, a feature commonly seen in apoptosing cells was not observed (Barros et al., 2003), which was a concern due to the exposure of neurons to the 1 minute pulse of H₂O₂.

We also observed that the labeling patterns as visualized by neutravidin staining (Fig. 2-5) was more diffuse than the immunostaining pattern for the epitope on the protein. Our hypothesis for this diffuse labeling was: 1) HRP is able to label neighboring cells possibly by generating enough radical to occupy all possible labeling sites (available tyrosine side chains) within the short labeling time 2) labeling is occurring on either the polylysine and laminin coated coverslips. Omission of laminin did not remove this diffuse labeling pattern, eliminating the latter possibility.
Figure 2-5. HRP is active on the neuron cell surface for BP. HRP is fused to the N terminus of a minimal transmembrane domain from PDGFR and introduced into neurons for lipofection. Cells were labeled live with 100uM BP in the presence of H$_2$O$_2$ as a co-oxidant for one minute, then fixed with 4% paraformaldehyde (PFA). A negative control is showed where H$_2$O$_2$ has been omitted. Biotinylation was detected by neutravidin fused to Alexafluor568. HRP was detected by immunostaining for the HA tag on it. Note that cells were not permeabilized.

Satisfied at least that HRP was active in neurons for biotin-phenol, we decided to focus on first tackling two logistical problems related to using HRP to label proteins in the synaptic cleft for proteomics: 1) Specific expression of HRP only within neurons, 2) identifying a transgene expression system that would allow us to express HRP in a large number of neurons needed to generate the amount of biotinylated material necessary for mass spectrometry.

Specific expression in neurons using a neuron specific promoter

One of the most commonly used promoters for expression of transgenes in mammalian cells is the CMV promoter, which was previously used to drive expression of APEX and HRP in HEK 293T cells used in the mitochondrial and IMS proteomes (Hung et al., 2014; Rhee et al., 2013). Because cultured cortical neurons (the model system we would apply our experiments in) are derived from embryonic brain tissue, they contain both neurons, and glial cells such as astrocytes, oligodendrocytes and other support cells, and it was necessary to identify a promoter that would only allow expression of HRP within neurons to prevent non-specific labeling of non-neuronal proteins not related to the synapse. The human synapsin promoter has previously been demonstrated to be specific for neuron expression(Hioki et al., 2007), and a side by side comparison of the
synapsin promoter and the CMV promoter expressing GFP to mark identify cell types in cultured cortical neurons showed neuron specific expression for the synapsin promoter, and a surprising preference for glial cells by the CMV promoter (Fig. 2-6). This convinced us to use the synapsin promoter to drive transgene expression for our experiments.

Figure 2-6. Synapsin promoter gives neuron specific GFP expression as identified by cell morphology in the DIC. The CMV promoter shown below drove GFP expression primarily in glial cells and not neurons.

**Transgene delivery of HRP to neurons**

To carry out a proteomics experiment, we would need a transfection system that was able to express HRP in a large number of neurons, so that we could label enough synaptic material with biotin to enrich and carry out our proteomics experiment. Simultaneously, the system should give us the flexibility to optimize the labeling and expression conditions in a large number of candidate HRP fusions to target HRP to the synaptic cleft. Lipofection and calcium phosphate transfection did not give a high transfection efficiency (~5% at best) when used on cultured cortical neurons, despite working well with fibroblasts (>80%) such as HEK 293T cells in the lab's previously published proteomes. These transfection methods often also resulted in very high expression levels that could not be tuned to a level that was low enough for synaptic labeling (discussed later) Since neurons are primary cells and post-mitotic, it was not possible to generate stable cell lines for HRP expression, while nucleofection
was found to generate a high propensity for cell death and required a large number of cells as starting material, while stipulating transgene expression throughout the entire lifetime of the neurons which were being cultured, which could lead to developmental artifacts since we were expecting to carry out our proteomics experiment at days in vitro (DIV) 18 and longer, when neurons were known to have developed and many mature synapses. Since generating transgenic mice for each HRP fusion to be tested would be prohibitively costly and time consuming, we decided to turn towards viral infection to express HRP, since expression levels and transfection efficiency could potentially be tuned by the multiplicity of infection (MOI), while viruses could be introduced late during neuronal development when mature synapses had already formed, reducing potential developmental artifacts, and viruses could be pseudotyped with the right cell surface proteins for neuronal infection. 3 viruses are commonly used to infect neurons, HSVs (Neve et al., 2005), lentiviruses (Naldini et al., 1996) and adenoviruses (Malik et al., 2012). We compared HSVs and lentiviruses because adenoviruses were restrictive in the size of the transgene that could be packaged and delivered (~4000bp), and many cell surface proteins that we were considering for delivery of HRP with were large and could not be delivered by adenovirus. We found that HSVs and lentiviruses both delivered the transgene efficiently into cultured cortical neurons with high infection efficiencies, but HSVs were more difficult to work with because these drove very strong expression of the transgene very quickly (<6 hours) after infection that was difficult to tune down for synaptic targeting (Fig. 2-7). Lentiviruses were however found to fit the criteria needed for our experiments; not perturbative in terms of morphology, expression was very low in general and could be tuned higher easily by changing the MOI, and generally easy to optimize since their generation was conducted within the lab and not generated by a core facility like HSVs were. This convenience and suitability drove us to use lentiviruses to deliver our constructs for the proteomic experiments.
Figure 2-7. HRP-NRX3b synapsin promoter HSV infected cortical neurons expression profile of HRP-NRX3b at (A) 6 hours and (B) 20 hours. These were compared with an IRES GFP HSV as a negative control for immunostaining. Cells were immunostained and imaged live for detection of the HA epitope. HSVs were produced by the MIT viral core facility courtesy of Rachel Neve.

**BxxP - cell impermeable probe for biotinylation**

For proteomic labeling, the use of our previously described biotin-phenol (BP) (Rhee et al., 2013) substrate would result in tagging of both cell surface and intracellular protein pools, proximal to HRP fusion constructs along the entire secretory pathway (Fig. 2-8), which would confound any analysis of the subsequently enriched proteins and their spatial localization. To restrict tagging to the surface and cleft space only, we sought to design a biotin-phenol variant that does not cross cell membranes. The cell impermeable probe must still be used as effectively by HRP as the original biotin-phenol probe and have similar chemical reactivity for tyrosine residues on the surface of endogenous proteins. In the application of APEX for mapping the mitochondrial matrix proteome (Rhee et al., 2013), a number of biotin-phenol derivatives were synthesized and evaluated for their ability to biotinylate endogenous proteins, both on the cell surface by HRP and within the cell by APEX (Fig. S2 of Rhee et al 2013).
typical criterion for generating a cell impermeable probe that cannot cross the phospholipid bilayer of the plasma membrane would be a negatively charged functional group or a derivative that is as large as glucose and also polar (Cooper, 2000). One candidate that caught our attention was the probe BxxP (Fig. 2-9). This probe has a polar, polyamide linker introduced between the phenol and biotin in BP to give BxxP (structures shown in Fig. 2-9). Previously, this probe was shown to have no activity with a cytosolic expressed APEX-NES (NES stands for nuclear excluded signal), while still maintaining robust activity with HRP expressed on the cell surface of HEK 293T cells. This suggested the probe might be cell impermeable, and so far we have not identified a biotin-phenol derivative that could be used by HRP and not APEX.

To test the hypothesis that BxxP was cell impermeable, we pre-incubated BxxP or BP with cultured cortical neurons expressing HRP expressed on the surface (HRP-NLG1), or in the cytosol (PSD95-APEX2 or APEX2-NES; NES is a nuclear exclusion signal) for 30 minutes (note that BxxP is normally concomitantly added together with 1mM H$_2$O$_2$ for 1 minute for labeling), before the addition of H$_2$O$_2$ to initiate biotinylation for 1 minute. Cells were then fixed,
permeabilized and stained for biotinylation using neutravidin-AlexaFluor 647 (Nv-647). Under these conditions, HRP at the cell surface gives biotinylation with both BxxP and BP, whereas intracellular peroxidase fusion constructs (PSD95-APEX2 and APEX2-NES) show biotinylation only with membrane-permeant BP. The absence of neutravidin staining with these constructs after BxxP treatment demonstrates that BxxP fails to enter neurons. Additionally, when the cell permeable BP was concomitantly added together with H2O2 for 1 minute for labeling, biotinylation was observed inside cells even within this short 1 minute window. The observation that BP was present in sufficient levels to be used by an intracellular APEX fusion (PSD95-APEX2) to give a biotinylation signal that was higher than background convinced us that it would be inappropriate for the synaptic cleft proteomics. To answer the question about whether APEX was able to use BxxP as efficiently as HRP (which is an alternative explanation that could explain the observations by Rhee and co-workers, APEX2 was also later shown by Dr. Phillip Stawski, a post-doctoral fellow in the Ting lab, to be able to use BxxP for biotin-phenol labeling on the cell surface with a number of surface expressed APEX constructs such as APEX2-GluA1 (data not shown), further supporting the hypothesis that BxxP is a cell impermeable probe.

Figure 2-9. Chemical structures of biotin probes used for proteomics. Structure of BxxP, which is membrane-impermeant, and BP, which is membrane permeant and used for intracellular proteomic mapping (Hung et al., 2014; Rhee et al., 2013).
Figure 2-10. Characterization of BxxP membrane impermeability. Cultured cortical neurons expressing one of the three indicated peroxidase fusion constructs at left were labeled with either BxxP (structure in 2-9) or membrane-permeant BP probe (Rhee et al., 2013). After 1 minute in the presence of H₂O₂, the neurons were fixed and stained with neutravidin-AlexaFlour647 to visualize biotinylation, and anti-V5 antibody to visualize peroxidase expression. HRP at the cell surface gives biotinylation with both BxxP and BP, whereas intracellular peroxidase fusion constructs (PSD95-APEX2 and APEX2-NES; NES is a nuclear exclusion signal) show biotinylation only with membrane-permeant BP. The absence of neutravidin staining with these constructs after BxxP treatment demonstrates that BxxP fails to enter neurons. Images are normalized except for the right column that shows the neutravidin-AlexaFlour647 signal at higher contrast.

HRP mediated biotinylation of synapses

The most crucial question that needed to be answered before we would know if the strategy described in Fig. 2-1 would succeed was if it was possible to actually label synapses specifically. This would entail 1) specific targeting of HRP to synapses, followed by 2) specific biotinylation of synapses.

Targeting of HRP to synapses

We rationalized that we could achieve HRP targeting to synapses by fusing HRP to the surface exposed domain of a synaptically targeted protein. We
fused HRP N terminus of a number of synaptic proteins, immediately after the signal-sequence, which is removed once the protein has been targeted to the secretory pathway for surface expression. These include the adhesion proteins neurexin 3beta (Nrx3b) and neuroligin1 (NLgn1), and the ion channel glutamate receptor subunit GluA2. We found that GluA2 did not tolerate the HRP fusion (as visualized by its poor surface targeting) and decided to focus on NLgn1 and Nrx3b, a well-studied synaptic adhesion complex that was expected to be present at excitatory glutamatergic synapses (Ichtchenko et al., 1996; Levinson et al., 2005).

These HRP fusions were cloned into our synapsin promoter lentiviral vector, and expressed in cultured cortical neurons. Markers for excitatory synapses such as the venus fluorescent protein (a GFP variant) fused to the C terminus of the post-synaptic density scaffolding protein Homer1b-venus, or the endogenous presynaptic marker Bassoon were visualized together with immunostaining of the surface population of HRP-NLGN1 or HRP-NRX3b, as visualized by the V5 or HA epitope tags on these constructs. Although targeting to synapses seemed promising for the neuroligin 1 construct even at the relatively high levels of expression by lipofection (Fig. 2-10), as viewed by the colocalization with the V5 staining (a non-synaptic pool was still observed here), the high levels of expression of the neurexin construct, even when introduced by lentiviral infection (Fig. 2-11), resulted in diffuse non-synaptic targeting.
Figure 2-11. HRP-NLG1 colocalization with Homer-Venus. Neurons were transfected using lipofectamine with FSW-HRP NLG1/ V5-NLG1 and Homer1b-venus and imaged at DIV 19. Neurons were immunostained live for the V5 tag on NLGN1 to visualize the surface pool of NLGN1, followed by fixation.

Figure 2-12. HRP-NRX3b colocalization with endogenous Bassoon. Neurons were transfected by lentiviruses with HRP-NRX3b or NRX3b and imaged at DIV 18. Live surface immunostaining of the HA tag on NRX3b, followed by fixation. Cells were permeabilized and immunostained for the endogenous marker Bassoon.

These results eventually made us decide to abandon our efforts with neurexin, and continue using neuroligin 1 as a platform for delivering HRP to synapses. At this point, we wanted to know if we could identify conditions for specific biotinylation of synapses, since all data that had been obtained so far
showed a diffuse labeling pattern that was larger than the normal size of synapses and also outside of synapses. Although our lab has previously found that tight, localized labeling of the subcellular structure one is trying to obtain the proteome of is not necessary for achieving a successful proteomics experiment (Hung et al., 2014), and a ratiometric tagging strategy (discussed later in this chapter and demonstrated first in (Hung et al., 2014)) could be used to differentiate between proteins that were in the structure versus outside of it. Nonetheless it was still difficult to assess whether; the proteins biotinylated were indeed all synaptic, if the diffuse labeling was biotinylation of extra-synaptic proteins, or if the diffuse labeling pattern was a product of proteins first labeled within the synapse, but diffused out of the synapse after they had been biotinylated. These questions prompted us to try to develop assays to improve the synaptic labeling of HRP-NLGN1, both by optimizing the expression levels and labeling conditions used.

Figure 2-13. Diffuse BxxP labeling from HRP fusions to NLGN1 and NRX3b.

Improving the labeling radius by co-quenching during labeling

We decided to explore the possibility that the labeling radius could be made tighter by the addition of radical quenchers during the 1 minute labeling window, a speculation made by Cronan and co-workers in their exploration of
pBirA’s labeling radius (Choi-Rhee et al., 2004). To assay for tighter labeling, we decided to explore a series of radical quenchers that are known to react with phenoxy radicals, of which a number (such as ascorbate, trolox, cysteamine) had previously been tested in HEK cells by a graduate student within the lab, Dr. Peng Zou. Phenoxy radicals can be quenched using small molecules that are able to donate a hydrogen to the radical, resulting in termination by reduction to a phenol. Two quenchers that had emerged as potential quenchers of the labeling reaction were gallate and mercaptoethanol, structures shown here (Fig. 2-14).

The labeling radius was assayed by using the following system: neurons were transfected with HRP-NLGN1 together with a venus fluorescent protein marker then fixed at DIV16 with 4% paraformaldehyde. Following fixation and washing with dPBS, cells would then be labeled for 1 min with BxxP and H2O2, in the presence of a different quencher in the concentrations as indicated (Fig. 2-15). Under these conditions, we could eliminate one of the potential sources for diffuse labeling; that proteins could be first labeled within the synapse, but diffused out of the synapse after they had been biotinylated. HRP has already been demonstrated to be active following fixation (these are the conditions normally used to generate DAB polymerization for electron microscopy (Martell et al., 2012).

Gallate and mercaptoethanol both reduced the labeling radius and intensity of labeling when used at 50nM and 100uM or 10uM respectively (Fig. 2-15, Fig. 2-16 shows how the analysis was performed). The 100uM mercaptoethanol condition seemed to reduce the labeling radius, while simultaneously did not reduce the labeling intensity significantly compared to an unquenched control (Fig. 2-15). We thought this might be a suitable choice for
decreasing the labeling radius and would need to explore how this looked when used in conjunction with live labeling of neurons.

Until this point, we had never been able to observe punctate, synapse like labeling that colocalized with the tips of dendritic spines (Fig. 2-15); these protrusions are known to receive synaptic input from axons (GRAY, 1959) and can be approximated as sites of synapses. This made these results very encouraging as they suggested that our system with optimization could potentially achieve specific labeling of synapses.

Figure 2-15. BxxP labeling in the presence of various radical quenchers with HRP NLG1 expressing neurons after post fixation gives varying intensities of labeling that are tight to diffuse. DIV 16 cortical neurons expressing HRP NLG1 and a cotransfection marker venus were first fixed, then labeled with BxxP for 1 min with 1mM H₂O₂ in Tyrodes buffer together with a quencher at the concentrations indicated. Neurons were then washed three times in Tyrodes then stained for biotin detection.
Figure 2-16. Analysis for BxxP labeling radius in the presence of quenchers. Smooth curve analysis was carried out across a line drawn perpendicular to the process. The intensities at each pixel along the line were first normalized from 0-1 using the pixel of highest intensity. The normalized intensities were then plotted using an X-Y plot. The distance (which can be obtained from the pixel position) between the half width of the expression channel (in blue) and the labeling channel (in red) was used to calculate an approximation of the labeling radius.

Low expression levels are sufficient for obtaining tight synaptic-like biotinylation patterns

While we explored conditions for reducing the labeling radius using chemical quenchers, we simultaneously explored the possibility that the expression levels of the HRP fusion might also influence the labeling radius observed. Since HRP NLGN1 was being introduced into neurons using a lentivirus, we began by titrating the amount of viruses used. To our pleasant surprise, we found that the lentiviral expression system under the synapsin promoter, was able to introduce our HRP-NLGN1 fusion at expression levels low enough that even in the absence of co-quenchers, it was possible to obtain biotinylation patterns from live cell biotinylation that looked very similar to when neurons expressing HRP-NLGN1 was first fixed, then labeled with BxxP as described in the co-quenching experiments above (Fig. 2-17). Note that at these expression levels, the epitope tag (V5) on HRP-NLGN1 could not be detected and only the biotinylation pattern was detectable.
Figure 2-17. Lentiviral infected neurons expressing HRP NLG1 can give BxxP labeling that resembles labeling of dendritic spine when labeled live (left panel showing BxxP detection by neutravidin 647), which is similar to lipofected neurons expressing the same construct which are labeled with BxxP post fixation (panels on right show the BxxP labeling channel as detected with neutravidin-647 and a venus cotransfection marker). Scale bars, 10 μm.

*HRP-NLGN1's biotinylation pattern colocalizes with synaptic markers*

Since we had found conditions that could give what was suggestive of synaptic labeling, we proceeded to characterize this labeling to validate that it was indeed synaptic. We wanted to ask two questions: 1) does the biotinylation pattern localize with a post-synaptic marker, such as Homer1b, and 2) is the biotinylation pattern preferentially localized to a particularly type of synapse? We first analyzed the localization of BxxP labeling with endogenous Homer1b protein, a post-synaptic scaffolding protein that is also a marker of glutamatergic synapses (Hayashi et al., 2009). BxxP labeling localized well with Homer1b puncta (Fig. 2-18), suggesting that most of the population of HRP-NLGN1 was localized to synapses.

Figure 2-18. Fluorescence imaging of HRP-NLGN1 with respect to synapse marker Homer. For maximal detection sensitivity, the HRP constructs are visualized via BxxP labeling for 1
minute followed by staining with neutravidin-AlexaFluor647 (red). Synapses (blue) are visualized with anti-Homer antibody. Scale bars, 10 μm.

Figure 2-19. Fluorescence imaging of HRP-NLGN1 with respect to excitatory and inhibitory synapse markers. HRP-NLGN1 is visualized via BxP labeling for 1 minute followed by staining with neutravidin-AlexaFluor647 (green). Excitatory synapses (red) and inhibitory synapses (orange) are visualized with anti-vGlut1 and anti-vGAT antibodies, respectively. Scale bars, 10 μm.

We then decided to explore the second question of what types of synapses are labeled by HRP-NLGN1 (Fig. 2-19). Using a similar colocalization analysis, we observed a similar distribution of HRP-NLGN1 labeled puncta colocalizing with vGlut and vGat as that observed for spines, suggesting that HRP-NLGN1 in neuronal culture might not be exceptionally specific for one type of synapse over the other although it definitely had a preference for excitatory synapses. This observation could be rationalized by a few hypotheses. One that HRP-NLGN1 has lower specificity for the type of synapses it is targeted to than the endogenous NLGN1, although similar values as what we observed for colocalization have previously been reported for the synapse specificity for NLGN1 (Chih et al., 2006). Another reason could be the non-specific staining of the antibody used, which could give some level of overlap between the vGlut and vGAT stain (a longer discussion of HRP NLGN1’s differential targeting between excitatory and inhibitory synapses is discussed in Chapter 3). Nonetheless, we believed that this construct was sufficiently competent for labeling synapses and proceeded to continue characterizing it.
Ultrastructure localization of HRP fusions in the synapse by electron microscopy

Utilizing HRP for electron microscopy in cultured neurons

As an alternate methodology for validating the synaptic targeting of HRP, we took advantage of the fact that HRP also serves as a GFP equivalent reporter for electron microscopy (Martell et al., 2012). To demonstrate that HRP, when fused to NLGN1 was competent and functional for DAB polymerization and also to optimize the labeling, staining conditions necessary for electron microscopy, we first expressed HRP in cultured cortical neurons and carried out DAB staining using a protocol similar to that adopted by Martell et al (Martell et al., 2012). Briefly, neurons expressing the HRP fusion were fixed first with glutaraldehyde at 37 degrees Celsius, then placed on ice and fixed for one hour. Cells were then stained with DAB followed by OsO₄, which oxidizes the DAB polymer and introduces osmium into the polymer to give contrast for electron microscopy. Samples were further stained with uranyl acetate, and then washed and embedded in resin before being sectioned for EM. These HRP fusions were expressed at relatively high levels since they were introduced by lipofection but our initial rationale for doing this was to determine if HRP has sufficient activity to give sufficient DAB staining to generate the contrast needed for EM. It was

Figure 2-20. Quantification of HRP-NLGN1 BxxP labeling's colocalization with endogenous synaptic markers. Quantitation of colocalization extent for images shown in Fig. 2-18 and 2-19 from >6 fields of view with >600 "puncta" Errors, ±1 s.d.
surprising however to see that the DAB staining for these constructs, especially when expressed at levels that were lower (synapsin promoter vs the CAG promoter), was very weak, and would likely be even weaker if we were to use lentiviruses to deliver HRP into these neurons (Fig. 2-21). We hypothesized that heme addition might improve the activity of HRP because heme availability might be limiting in neurons, and Dr. Jeffery Martell, another graduate student in the lab had previously found that addition of heme would increase peroxidase (APEX) activity in HEK 293T cells. Addition of heme the day before fixation dramatically increased the activity of HRP, giving much darker DAB contrast (Fig. 2-21), even when expressed at low levels using lentiviral infection and we decided to use these conditions to evaluate the submicron level synaptic targeting of the HRP fusions.

![Figure 2-21](image)

**Figure 2-21.** DAB staining (20min) for DIV 20 hippocampal neurons transfected w/ HRP fusions by L2K for 2 days and incubated with 7uM heme for 16 hours. DAB staining was carried out for 20 min after fixation with glutaraldehyde in sodium cacodylate buffer.

*Synaptic localization of HRP-NLGN1 and HRP-NRX3b*

Confident that we had sufficient peroxidase activity to detect DAB staining by electron microscopy, we infected neurons with HRP-NLGN1 and HRP-NRX3b to evaluate their sub-micron targeting to synapses. These were introduced into cultured cortical neurons at DIV15, and fixed on ice at DIV19 in preparation for DAB staining with glutaraldehyde. At the light microscopy level, we observed clear DAB staining of the cell body (soma) of the HRP expressing neurons as visualized by bright field light microscopy, which reflects the extent at which the
sample absorbs light illuminated on it, allowing us to identify regions stained strongly by the DAB polymer. At the EM level, staining was observed restricted to the extracellular space, between the plasma membranes of two cells, and almost always juxtaposed by synaptic vesicles on one side of the membrane, suggesting that these regions were indeed synapses. For HRP-NRX3b, trafficking vesicles which were stained dark were also observed, in the vicinity of synaptic vesicles suggesting that this might be a presynaptic terminus. These stained vesicles could perhaps represent HRP-NRX3b being delivered to the presynaptic terminus. The staining pattern for both HRP-NLG1 and HRP-NRX3b was both non-uniform but periodic within the synaptic cleft, and we observed a toothlike pattern within the presynaptic side that seemed to localize with the dark stain in the cleft. Such toothlike patterns have been observed by electron microscopy before, and were speculated to be active zone densities, protein dense structures that prime and prepare synaptic vesicles for release into the cleft similar to the T-bars observed in Drosophila synapses (Südhof, 2012). Because the DAB polymer does not cross membranes, these “teeth” are unlikely to be contributed by HRP activity within the cleft leaking into the presynaptic space.

HRP NLGN1

HRP NRX3b
Figure 2-22. HRP-NLGN1 and HRP-NRX3b are targeted to synapses as visualized by electron microscopy. Constructs were introduced into DIV19 neurons at DIV15, and HRP catalyzes the polymerization and local deposition of diaminobenzidine, which in turn recruits electron dense osmium (Martell et al., 2012). The dark-stained regions in the synaptic clefts indicate the presence of HRP. Panels on the top are imaged with bright field microscopy using a 40X microscope. Panels below are the matched samples imaged with electron microscopy by Tom Deerinck and Mark Ellisman at UC San Diego.

**Streptavidin bead enrichment of synapses for proteomics**

Confident that we could now target HRP to synapses using HRP-NLGN1, and also biotinylate them as well, we decided to continue characterizing the biotinylation of proteins by the different HRP baits, and also the proteins that would be enriched by streptavidin bead enrichment, towards the eventual goal of generating sufficient biotinylated material from synapses for proteomic identification. We first assayed for biotinylation using a streptavidin-HRP conjugate western blot of neuronal whole cell lysates generated from cortical neurons that were infected with the three different HRP baits using lentiviruses (HRP-NLGN1, NRX3b and TM), and labeled with the cell impermeable BxxP or the cell permeable BP (Fig. 2-23). APEX-NES was also included as a control for intracellular labeling, and uninfected neurons were included as a negative labeling control. We observed strong biotinylation of endogenous proteins that was HRP dependent when BxxP was used, and also a different labeling pattern of proteins between HRP expressed generally on the cell surface, compared to HRP targeted to synapses (HRP-NLGN1 and HRP-NRX3b). Intracellular APEX-
NES did not show any significant labeling as compared to the negative labeling control (the endogenous biotinylated proteins are visible in this blot), further supporting the hypothesis that BxxP is cell impermeable. A different result is obtained when the cell permeable BP is used, which showed biotinylation of endogenous proteins even in the APEX-NES infected cells (note that this was using the first generation of APEX, which in general shows much less activity than HRP in neurons).

Figure 2-23 HRP-NLGN1 gives a distinct biotinylation pattern of proteins as visualized by streptavidin HRP western blot. Whole cell lysates of 0.5 million DIV 18 cortical neurons infected with lentiviruses expressing HRP fusions were labeled with biotin-xx-phenol, quenched and lysed and run for streptavidin HRP westerns. Intracellular labeling was also carried out of similarly infected neurons, but with biotin-phenol instead and with 30 min preincubation. A red arrow indicates the endogenous biotinylated protein band.
Since we were able to biotinylated endogenous proteins other than the HRP baits themselves, we proceeded to carry out enrichment of the biotinylated proteins using streptavidin magnetic beads. Streptavidin bead enriched eluates showed strong and unique biotinylation protein patterns when HRP-NLGN1 and HRP-TM were present, while negative controls (shown in lanes 2, 4, and 5 in Fig. 2-24) showed only enrichment of the endogenous biotinylated proteins (Moss and Lane, 1971; Wood and Barden, 1977). We also found that at 7.5 million neurons used as starting material, the amount of pulled down material was enough to visualize using coomaissie, and was visible by silver staining which is more sensitive (Fig. 2-25).

![Figure 2-24. Streptavidin blot analysis of streptavidin-enriched lysates after live-neuron biotinylation with HRP NLGN1 and HRP-TM constructs. 13 million DIV 18 cortical neurons infected with lentiviruses expressing HRP fusions (to cleft or non-specific cell surface) were labeled for 1 min with H2O2 and biotin-xx-phenol, quenched and lysed, then pulled down using streptavidin magnetic beads overnight (50ul per condition). Proteins were eluted in 50ul of loading buffer, 2mM biotin, 20mM DTT. Eluate was run on 8% SDS PAGE and visualized by blotting with streptavidin-HRP (left) Numbers correspond to the following conditions: 1: HRP NLG1+BxxP+ H2O2, 2: 1 but omit HRP, 3: HRP-TM+BxxP+ H2O2, 4: 1 but omit H2O2, 5: 1 but omit BxxP. Blot on the right is the same blot on the left but exposed for a longer period of time.](image-url)
Preliminary synaptic cleftome using Neuroligin 1

The following section describes our first attempt at mapping the synaptic cleft proteome. This experiment provided us with a preliminary dataset to identify ways to improve our proteomic mapping methodology, and subsequent analysis of the dataset itself. Because many of these sections are repetitive with the actual successful mapping of the synaptic cleft proteome described later in this chapter, the following sections will focus on the experimental design and potential improvements that were learnt and implemented for the final proteomic mapping of the synaptic clefts experiment.

Synaptic Cleftome using HRP-NLGN1 and iTRAQ for quantitative proteomics

Since we were able to demonstrate that HRP-NLGN1 was synaptically localized (by electron microscopy and fluorescence microscopy), able to biotinylate synapses, and we were able to enrich using streptavidin beads sufficient biotinylated material, we proceeded to focus on designing a proteomics experiment using HRP-NLGN1. Contrary to previous proteomic attempts involving APEX, we did not use a SILAC (Stable isotope labeling by amino acids in cell culture (Ong et al., 2002) based proteomic approach because neurons are primary cells and do not divide, hence metabolic incorporation of isotopically
labeled amino acids would not be comprehensive of the entire proteome. As a result, this required a different quantitative proteomics strategy; we employed a strategy called iTRAQ that is compatible with non-dividing cells, where chemical labeling of the isotopic tags can be implemented following cell lysis.

**iTRAQ quantitative proteomics**

iTRAQ or Isobaric tags for relative and absolute quantitation (iTRAQ) is a labeling method used in quantitative proteomics in conjunction with tandem mass spectrometry to quantify the relative abundances of protein from different sources (Ross et al., 2004). For the purpose of this thesis, we focused on up to 4plex or 4 states per mass spectrometry experiment although iTRAQ is also able to achieve up to 8plex. In iTRAQ, isobaric (same mass) reagents are used to label the primary amines of peptides and proteins, using a N-hydroxy succinimide ester group that is reactive for the N terminus of peptides (Fig. 2-26). Each reagent also contains a unique reporter group and its respective balance group, which together make the labeled peptides generated from each state isobaric. This is possible because the reporter groups have masses of 114, 115, 116 or 117, and are combined with a balance group of 31, 30, 29 or 28 respectively.
Protein samples cultured under different conditions (as indicated by the different colors) are harvested for proteolytic digestion and labeled with unique iTRAQ reagents (structure shown in bottom right) that label the N-terminus of peptides generated. These are then mixed together for LC-MS/MS identification. In the mass spectrometer, the backbone fragmentation pattern identifies the peptide sequence, while the reporter ion intensities identifies the relative abundance between different states.

The iTRAQ workflow is as follows (Fig. 2-26): samples are first prepared following under the different states in which they would like to be assayed. Cells are lysed to extract proteins, and the input protein is normalized between states using a standard protein assay such as a Bradford assay. Once normalized, proteins can then be digested using a protease such as trypsin, to generate proteolytic peptides that can then be labeled with the iTRAQ reagents at their N termini. The labeled peptides are then combined, and run together in a LC-MS/MS (liquid chromatography, mass-spectrometer/mass-spectrometer) setup, that functions at the MS/MS level as both a peptide sequencer based on the peptide backbone fragment ions, and as a way to quantify between states by comparing the fragmented reporter ions. The main difference between SILAC and iTRAQ as quantitative methods are that the peptides are labeled following trypsin digestion and then combined for running in the mass spectrometer which is different from SILAC where the mixing occurs at the protein level. This can inherently introduce errors that come from handling steps occurring before the
mixing at the peptide level, but is a small price to pay if the cells themselves cannot metabolically incorporate the labeled amino acids. We chose to use iTRAQ because we encountered inconsistencies involving efficient SILAC incorporation in primary neurons (also discussed in Chapter 5), which requires successive cell divisions to efficiently incorporate the isotopically labeled amino acids into the cell's proteome or long term culturing together with sufficient protein turnover to get efficient incorporation into the proteome.

**iTRAQ experimental design for synaptic cleft proteomics**

Our first attempt at the synaptic cleftome involved the following experimental set up (Figure 2-27). In experimental replicate 1, the first set of neurons would express HRP-NLGN1 to label synapses, concurrently with a second state that is a negative labeling control (omitting BxxP probe during the labeling step), and lastly a state that contains HRP fused to the TM domain of PDGFR (HRP-TM) which serves as a way to differentiate between general cell surface proteins, and proteins actually found within the synaptic cleft (a ratiometric tagging strategy developed in (Hung et al., 2014) for defining the proteomes of open compartments). Experimental replicate 2 contained a repeat of the HRP-NLGN1 synaptic cleft labeling state, and two negative control states; omission of HRP (no transfection) and omission of H$_2$O$_2$ during the labeling step. Each state was labeled with the iTRAQ reagents (115,116,117) described in Figure 2-27.

![Figure 2-27. Design of 2 independent proteomics experiments using HRP-NLGN1 and iTRAQ labeling. Each experiment consisted of three samples, which were separately enriched and tagged with unique iTRAQ labels.]
We cultured rat cortical neurons till DIV 14, at which cells were infected with lentiviruses that would express HRP-NLGN1, HRP-TM or a mock infection (using neurobasal media in which the viruses are diluted in). The viruses used in these experiments, were titered to levels at which they gave punctate “synaptic” like labeling as seen in Figure 2-17 and assayed for the various synaptic markers (Fig. 2-18-19), and were also observed at synapses at the sub-micron level by electron microscopy (Fig. 2-22). Neurons were labeled with BxxP and H₂O₂ at DIV 18 (or with the appropriate negative control for those states). A paired imaging experiment that was concurrently labeled is presented in Figure 2-28. Neurons were then lysed with RIPA buffer, normalized for protein input (1.5mg of protein was used per state which corresponded to 3 10cm dishes of neurons, 7.5 million neurons in total) and 150ul of streptavidin magnetic beads per state was used to enrich for biotinylated proteins by overnight incubation at 4°C. The next day, beads were washed using 2 washes with RIPA buffer, a high salt 1M KCl wash, followed by a high pH 0.1M Na₂CO₃ (pH 10.5), a 2M urea wash and then another 2 washes with RIPA. These washes were found to be necessary for eliminating non-specific binders and improving the eventual proteomic data quality (development of these conditions is discussed further in Chapter 5). After washing the beads were turned over to our proteomics collaborators Dr. Udeshi and Dr. Svikina in Dr. Steven Carr’s laboratory at the Broad Institute. Proteins that had been enriched were then digested with trypsin on beads to generate peptides that could be labeled with the respective iTRAQ reagents. These labeled peptides were then mixed and run together in the LC-MS/MS using a Q-exactive instrument.
Figure 2-28. Imaging the biotinylation activities of HRP fusion constructs used for the first synaptic cleft proteomics experiment. Fluorescence imaging of the biotinylation catalyzed by HRP fusion constructs used for proteomics. DIV19 rat cortical neurons expressing the constructs indicated across the top were treated with BxxP and H$_2$O$_2$ for 1 minute, then fixed and stained with neutravidin-AlexaFluor647 to detect biotinylated proteins (middle row) and anti-V5 antibody to detect HRP construct expression (top row). The second and last two columns show negative controls with HRP, H$_2$O$_2$ or BxxP omitted respectively. The bottom row shows an overlay of all 3 channels with the DIC.

Characteristics of the proteomic data

The mass spectrometry data was searched using the Rat Uniprot database with the Maxquant analysis program (Cox and Mann, 2008). Following searching of the dataset we identified 770 proteins that had 2 or more unique peptides identifying that protein in both experimental replicates, a standard for proteomics, this was a relatively low number (compared to >2000 proteins for the mitochondrial matrix proteomes), possibly due to the low amount of protein input used. Of these 70 were known synaptic cleft proteins (out of 158 proteins from a true positive list comprising known synaptic cleft proteins). Promisingly however, the dataset showed that we were enriching for surface proteins as shown in Fig. 2-29, and we were also deenriching for mitochondrial proteins, which was a little surprising and we hypothesize that this occurs because endogenous biotinylated proteins are preferentially enriched in the negative controls, but surprising that
many other non-endogenous biotinylated proteins in the mitochondria were also deenriched.

Figure 2-29. Cell surface proteins are enriched in the proteomic data generated using HRP-NLGN1. Scatterplot of all proteins within the proteomic dataset, with the x-axis representing HRP-NLGN1/omit HRP, and the y-axis the ratio for HRP-NLGN1/omit H2O2. Mitochondrial proteins are marked in orange, and known surface proteins are marked in blue.

Synaptic proteins are enriched in our synaptic cleftome.

To provide ourselves with more confidence that we were indeed enriching synaptic proteins and not just surface proteins we plotted histograms of the known true positive proteins (proteins found in the synapse with cleft exposure) and compared this to non-synaptic proteins such as mitochondrial and nuclear proteins (Fig. 2-30). True positives were enriched as compared to true negatives, although the distributions were not as bimodal as seen in the mitochondrial matrix proteomes (Rhee et al., 2013) suggesting that improvements could be made to improve the separation and enrichment of synaptic proteins.

Nonetheless, this was encouraging because it suggests that the experiment was working as envisioned and we had previously encountered a large number of non-specific binders observed in a previous proteomic attempt (the AIS proteome discussed in Chapter 5) and the bead washing conditions developed as a result
allowed for a successful enrichment of synaptic cleft proteins. Many of the most enriched proteins included glutamate receptors, such as the GluA subunits and NMDA receptors, and also structural components of the synapse such as the neuroligins. Neuroligin1, our vehicle for delivering HRP to the synapse was strongly enriched in the experiment, a good indicator that the enrichment process was working.

Because enrichment for surface proteins does not equate to synaptic enrichment since synaptic proteins are a subset of all surface proteins and we are confident of enriching those, we performed two analyses, one where we plotted a histogram of the enrichment ratios of known synaptic proteins and the non-synaptically annotated surface proteins found in our dataset (note that there may indeed be novel synaptic proteins found within this list). This showed a slight enrichment for synaptic proteins (Figure 2-31). Promisingly, a histogram of the ratio of synaptic proteins to the total number of surface proteins plotted against the enrichment ratio, shows a positive trend between the ratio of synaptic/general surface proteins and the enrichment ratio suggesting we are indeed enriching for synaptic proteins over other surface proteins (Figure 2-32).

Figure 2-30. Synaptic proteins are enriched in the proteomic data. Enrichment ratio histograms for true positives in green (synaptic proteins) plotted against false positives in red shows that synaptic proteins are enriched relative to the intracellular proteins.
Figure 2-31 Synaptic proteins are enriched slightly compared to general surface proteins in the proteomic data. Ratiometric tagging comparison of HRP-NLG1 enriched proteins vs HRP-TM shows a slight enrichment of synaptic proteins compared to non-annotated surface proteins, which may also include undiscovered synaptic cleft proteins.

Figure 2-32. Ratio of known synaptic proteins/total surface proteins increases as enrichment ratio (HRP NLGN1/HRP-TM) increases.

Cytoplasmic PSD proteins are also enriched

We next evaluated the ratiometric tagging strategy (Hung et al., 2014) by plotting a scatterplot of enrichment by HRP-NLG1 compared to HRP-TM (Fig.
By this strategy, synaptic cleft proteins should be enriched to a greater extent by HRP-NLGN1 as compared to HRP-TM, and hence lie closer to the HRP-NLGN1 axis (the y axis in this figure). The scatterplot did not suggest that it was easy to discern if all known synaptic proteins had a strong bias towards having a high NLG1/TM ratio (Figure 2-33), and we can only speculate this may reflect certain synaptic proteins having different amounts of extra-synaptic pools.

One thing we did notice however was that there were many cytoplasmic PSD proteins that had very high NLG1/TM ratios (shown as blue dots in the scatterplot). This was a little hard to reconcile, but the hypothesis we have for this is that our ratiometric tagging experiment is indeed succeeding, because HRP-TM cannot label cytoplasmic synaptic proteins (because BxxP cannot cross membranes) hence giving these proteins a low TM/negative control ratio (x axis in the figure), yet for these proteins to have a high HRP-NLGN1/negative control ratio, they would have to either labeled by HRP-NLGN1 (which would not be possible if BxxP is cell impermeable) or be pulled down via their direct interaction with synaptic proteins that are being labeled and we believe this might be the reason for this observation.

Because we know that the PSD is detergent resistant and a dense crosslinked mesh of proteins, we believe that under our current wash conditions, it is not sufficiently dissociated into its distinct protein components, resulting in complexes versus individual proteins being enriched. Prior literature shows it is possible to immunopurify the PSD and isolate more than 100 known PSD components via just antibodies to the NMDA ion channel subunit GluN1 (Husi et al., 2000) or scaffold protein PSD95 (Fernández et al., 2009) under similar immunopurification conditions (in RIPA buffer). This would need to be addressed to define an accurate synaptic cleft proteome. We can also rule out the possibility of the BxxP crossing cell membranes and labeling intracellular proteins, because the enriched cytoplasmic proteins are not ER/golgi proteins (where HRP is present and should directly label for enrichment if BxxP is cell permeable) but instead non-membrane bound cytoplasmic proteins, and we know from our prior experiments that the probe does not cross membranes.
Proteins in this region likely have a high synapse localization and no extra synaptic pool.

Proteins in this region likely have a high synapse localization and an extra synaptic pool.

Log$_2$ (HRP NLG1/omit BXXP ratio)

Log$_2$ (HRP TM/omit BXXP ratio)

Figure 2-33. Scatterplot illustrating synaptic proteins (green) are enriched compared to intracellular proteins (red). Cytoplasmic PSD proteins (blue) are also strongly enriched by HRP-NLGN1 (y axis) compared to HRP-TM (x axis).
<table>
<thead>
<tr>
<th>Protein</th>
<th>Abbreviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>IQ motif and SEC7 domain-containing protein 2</td>
<td>Iqsec2</td>
</tr>
<tr>
<td>Disks large homolog 3</td>
<td>Dlh3</td>
</tr>
<tr>
<td>Isoform 2 of Disks large homolog 2</td>
<td>Dlh2</td>
</tr>
<tr>
<td>Disks large-associated protein 1</td>
<td>Dlgap1</td>
</tr>
<tr>
<td>Protein Iqsec2</td>
<td>Iqsec2</td>
</tr>
<tr>
<td>Isoform 4 of Ras/Rap GTPase-activating protein SynGAP</td>
<td>SynGAP</td>
</tr>
<tr>
<td>SH3 and multiple ankyrin repeat domains protein 1</td>
<td>Shank1</td>
</tr>
<tr>
<td>Calcium/calmodulin-dependent protein kinase type II subunit alpha</td>
<td>Camk2a</td>
</tr>
<tr>
<td>Isoform Delta 6 of Calcium/calmodulin-dependent protein kinase type II subunit delta</td>
<td>Camk2d</td>
</tr>
<tr>
<td>Ras/Rap GTPase-activating protein SynGAP</td>
<td>SynGAP</td>
</tr>
<tr>
<td>Isoform 4 of SH3 and multiple ankyrin repeat domains protein 2</td>
<td>Shank2</td>
</tr>
<tr>
<td>Calcium/calmodulin-dependent protein kinase II, beta, isoform CRA-a</td>
<td>Camk2b</td>
</tr>
<tr>
<td>Disks large homolog 4</td>
<td>Dlh4</td>
</tr>
<tr>
<td>Brain-specific angioeisn inhibitor 3 (Predicted)</td>
<td>Bsi3</td>
</tr>
<tr>
<td>Protein Bal3</td>
<td>B3</td>
</tr>
<tr>
<td>Acid-sensing ion channel 1</td>
<td>AscC1</td>
</tr>
<tr>
<td>Calcium/calmodulin-dependent protein kinase type II subunit gamma</td>
<td>Camk2g</td>
</tr>
<tr>
<td>Noelin-2</td>
<td>Offm2</td>
</tr>
<tr>
<td>Gamma-aminobutyric acid receptor subunit beta-1</td>
<td>Gabrb1</td>
</tr>
<tr>
<td>Ankyrin repeat and sterile alpha motif domain-containing protein 1B</td>
<td>Anks1b</td>
</tr>
<tr>
<td>Receptor tyrosine-protein kinase erbB-4</td>
<td>ErbB4</td>
</tr>
<tr>
<td>Brain-specific angioeisn inhibitor 2 (Predicted)</td>
<td>Bal2</td>
</tr>
<tr>
<td>Gephyrin</td>
<td>Gphn</td>
</tr>
<tr>
<td>Protein Iqsec1, IQ motif and SEC7 domain-containing protein 1</td>
<td>Iqsec1</td>
</tr>
<tr>
<td>Inward rectifier potassium channel 4</td>
<td>Kir4b</td>
</tr>
<tr>
<td>Protein Pknsa5</td>
<td>Pknsa5</td>
</tr>
<tr>
<td>Membrane-associated guanylate kinase, WW and PDZ domain-containing protein 2</td>
<td>Magl2</td>
</tr>
<tr>
<td>Disks large-associated protein 3</td>
<td>Dlgap3</td>
</tr>
<tr>
<td>Connector enhancer of kinase suppressor of ras 2</td>
<td>Cks2</td>
</tr>
<tr>
<td>Gamma-aminobutyric acid receptor subunit alpha-3</td>
<td>Gabra3</td>
</tr>
<tr>
<td>Protein Efin1</td>
<td>Ef1n1</td>
</tr>
<tr>
<td>Noelin</td>
<td>Offm1</td>
</tr>
<tr>
<td>Protein Hecew</td>
<td>Hecew1</td>
</tr>
<tr>
<td>Protein gamma-aminobutyric acid receptor alpha-3</td>
<td>Gabra3</td>
</tr>
<tr>
<td>Protein Dlg5</td>
<td>Dlg5</td>
</tr>
<tr>
<td>Serine/threonine-protein kinase DCLK1</td>
<td>Dclk1</td>
</tr>
<tr>
<td>Protein Tanc2</td>
<td>Tanc2</td>
</tr>
<tr>
<td>Protein R601311344</td>
<td>Dsk1</td>
</tr>
<tr>
<td>Protein Sest1</td>
<td>Sest1</td>
</tr>
<tr>
<td>Disks large homolog 1</td>
<td>Dlh1</td>
</tr>
<tr>
<td>Disks large-associated protein 1 type 8 receptor subunit 1</td>
<td>Dlgap1</td>
</tr>
</tbody>
</table>

**Figure 2-34. Cytoplasmic PSD proteins are also enriched in this proteome.** Proteins are arranged from top to bottom by increasing NLG1/TM ratio. Cytoplasmic PSD proteins are highlighted in beige. Synaptic cleft proteins are highlighted in green.

**Limitations of this proteomic dataset**

One of the surprising results when using neuroligin1 as the HRP bait, was the enrichment of inhibitory synapse components, such as neuroligin2 and GABA<sub>A</sub> receptor subunits in addition to excitatory synapse components (Fig. 2-35). This matched our in vitro data describing that Neuroligin 1 was not complete excitatory synapse specific, as shown by our colocalization data. This would not
bode well if we were trying to identify specifically the synaptic cleftome of excitatory vs inhibitory synapses, and we would need to work on identifying synaptic proteins that could do this. From our dataset, it was clear that several improvements needed to be addressed for the next proteomic experiment. These were 1) eliminating cytoplasmic PSD proteins, 2) Obtaining additional synaptic targeted HRP fusions to generate an intersection of proteins for higher specificity 3) Obtaining synaptic targeted HRP fusions that could differentiate excitatory vs inhibitory synapses. We speculate that eliminating the cytoplasmic PSD proteins, which are in greater abundance that most synaptic cleft exposed proteins (Cheng et al., 2006) would improve the sensitivity of detecting synaptic cleft proteins, and improve the separation between synaptic cleft proteins and the cytoplasmic false positives. HRP fusions that are specific for each type of synaptic cleft would also make it easier to identify the function significance of proteins enriched by these HRP baits, since the two synaptic clefts have divergent and opposite functions (excitation vs inhibition).
<table>
<thead>
<tr>
<th>Protein name</th>
<th>Gene Symbol</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interleukin-1 receptor accessory protein-like 1</td>
<td>Ilirap1</td>
</tr>
<tr>
<td>Glutamate receptor 2</td>
<td>Gria2</td>
</tr>
<tr>
<td>Isoform Flp of Glutamate receptor 2</td>
<td>Gria2</td>
</tr>
<tr>
<td>Glutamate receptor ionotropic, kainate 3</td>
<td>Grik3</td>
</tr>
<tr>
<td>Neurogin-2</td>
<td>Nlgn2</td>
</tr>
<tr>
<td>Glutamate receptor ionotropic, NMDA 2B</td>
<td>Grin2b</td>
</tr>
<tr>
<td>Neurogin-1</td>
<td>Nlgn1</td>
</tr>
<tr>
<td>Isoform B of Glutamate receptor ionotropic, NMDA 1</td>
<td>Grin1</td>
</tr>
<tr>
<td>Isoform 2 of Glutamate receptor 4</td>
<td>Grin4</td>
</tr>
<tr>
<td>Glutamate receptor ionotropic, NMDA 2A</td>
<td>Grin2a</td>
</tr>
<tr>
<td>Brain-specific angiogenesis inhibitor 3</td>
<td>Bai3</td>
</tr>
<tr>
<td>Protein Bai1</td>
<td>Asic1</td>
</tr>
<tr>
<td>Acid-sensing ion channel 1</td>
<td>Neto2</td>
</tr>
<tr>
<td>Neurophin and tolloid-like protein 2</td>
<td>Olfm2</td>
</tr>
<tr>
<td>Noelin-2</td>
<td>Cacna1e</td>
</tr>
<tr>
<td>Voltage-dependent R-type calcium channel subunit alpha-1E</td>
<td>Cdh9</td>
</tr>
<tr>
<td>Protein Cdh9</td>
<td>Gabrb1</td>
</tr>
<tr>
<td>Gabrb1</td>
<td>Lrcc7</td>
</tr>
<tr>
<td>Gabra5</td>
<td>Gsn</td>
</tr>
<tr>
<td>Gaba5</td>
<td>Grin2b</td>
</tr>
<tr>
<td>Protein Copa</td>
<td>Cgsn</td>
</tr>
<tr>
<td>Brain-specific angiogenesis inhibitor 2</td>
<td>Bai2</td>
</tr>
<tr>
<td>Isoform Flp of Glutamate receptor 1</td>
<td>Gria1</td>
</tr>
<tr>
<td>Gabra5</td>
<td>Lrcc3</td>
</tr>
<tr>
<td>Gamma-aminobutyric acid receptor subunit alpha-5</td>
<td>Gria3</td>
</tr>
<tr>
<td>Gabra5</td>
<td>Gabra2</td>
</tr>
<tr>
<td>Gamma-aminobutyric acid receptor subunit alpha-2</td>
<td>Konj1</td>
</tr>
<tr>
<td>Inward rectifier potassium channel 4</td>
<td>Pdhb</td>
</tr>
<tr>
<td>Protein disulfide-isomerase</td>
<td>Dscam</td>
</tr>
<tr>
<td>Down syndrome cell adhesion molecule homolog</td>
<td>Lrcc4c</td>
</tr>
<tr>
<td>Leucine rich repeat containing 4C</td>
<td>Gria3</td>
</tr>
<tr>
<td>Isoform Flp of Glutamate receptor 3</td>
<td>Gabrb3</td>
</tr>
<tr>
<td>Gamma-aminobutyric acid receptor subunit beta-3</td>
<td>ELFN1</td>
</tr>
<tr>
<td>Protein Elfn1</td>
<td>Noelin</td>
</tr>
<tr>
<td>Noelin</td>
<td>Olfm1</td>
</tr>
<tr>
<td>Protein Ptprd (Fragment)</td>
<td>Ptprd</td>
</tr>
<tr>
<td>Solute carrier family 12 member 5</td>
<td>Slc12a5</td>
</tr>
<tr>
<td>Protein Sezil2</td>
<td>Sezil2</td>
</tr>
<tr>
<td>Protein Igsf9b</td>
<td>Igsf9b</td>
</tr>
<tr>
<td>Metabotropic glutamate receptor 4</td>
<td>Grm4</td>
</tr>
<tr>
<td>Receptor-type tyrosine-protein phosphatase S</td>
<td>Ptpns</td>
</tr>
<tr>
<td>Leucine-rich repeat-containing protein 4</td>
<td>Lrcc4</td>
</tr>
<tr>
<td>Gamma-aminobutyric acid receptor subunit alpha-3</td>
<td>Gaba3</td>
</tr>
<tr>
<td>Pultitary adenylate cyclase-activating polypeptide type 1 receptor</td>
<td>Adcyap1r1</td>
</tr>
<tr>
<td>Gamma-aminobutyric acid receptor subunit beta-2</td>
<td>Gabrb2</td>
</tr>
<tr>
<td>Protein Fam17a2</td>
<td>Fam17a2</td>
</tr>
<tr>
<td>Netrin G2</td>
<td>Ntng2</td>
</tr>
<tr>
<td>Interleukin 1 receptor accessory protein b</td>
<td>Ilyrap</td>
</tr>
<tr>
<td>Leucine-rich repeat-containing protein 4B</td>
<td>Lrcc4b</td>
</tr>
<tr>
<td>Glutamate receptor ionotropic, kainate 5</td>
<td>Grik5</td>
</tr>
<tr>
<td>Syntaxin-1B</td>
<td>Stx1b</td>
</tr>
<tr>
<td>Latrophilin-1</td>
<td>Lphn1</td>
</tr>
</tbody>
</table>
this procedure failed to remove intracellular proteins of the excitatory post-
synaptic density (PSD) which we hypothesized to be tightly crosslinked in a
detergent-insoluble meshwork (Carlin et al., 1980) to cleft-exposed
transmembrane proteins, resulting in their enrichment together with biotinylated
cleft exposed proteins (Fig. 2-35). The PSD itself, as explained in Chapter 1 is
detergent resistant, a chemical property that has been enlisted in its purification
from synaptosomes by extraction with triton-X 100 suggesting that detergent
concentrations in our original enrichment protocol (in the RIPA buffer) do not
contribute sufficiently towards separating protein-protein interactors. Further
support for this hypothesis were that previous attempts at isolating the PSD by
affinity purification, had found that they could enrich 118 synaptic proteins within
the PSD complex by pulling down PSD95, which constitutes one of the major
components of the PSD (Fernández et al., 2009), suggesting that the PSD itself
is tightly linked together in a mesh framework that can be isolated via a bait
protein.

Figure. 2-36. Rationale for an enrichment protocol that dissociates the PSD before
streptavidin bead enrichment of biotinylated proteins. Because the PSD is a stable
crosslinked complex, secondary interaction partners that are not biotinylated by HRP are
enriched under insufficiently denaturing conditions for streptavidin bead enrichment.

To eliminate cytoplasmic PSD proteins, we hypothesized that by first
breaking up the PSD, and then subsequently performing the pull-down in concert
with harsh denaturing washes, we should be able to eliminate these protein-
protein interactions. We screened a number of conditions to accomplish this, and
found that this could be achieved by lysis of the cell pellet in 1% SDS (sodium
dodecyl sulfate, a detergent normally added to lysis buffers, at a 10 fold concentration normally used in RIPA buffer, a buffer often used in affinity purifications) followed by boiling for 10 minutes. Although the boiling process denatures proteins and might be detrimental towards typical affinity purification pulldowns involving antibodies, we expect that this will not perturb the recognition of the biotin moiety by streptavidin.

After boiling, the lysate is subsequently diluted to the same concentration as RIPA buffer to generate a pulldown compatible buffer, before streptavidin magnetic beads are added to it for the pulldown step. This extra step helped to eliminate the retention of PSD95 a known component of the PSD complex and direct interactor of Neuroligin1 in the pulldown sample, but did not adversely affect the pulldown of synaptic proteins such as the glutamate receptor subunit GluA1 nor the HRP bait itself (compare conditions 1 and 2 in Fig. 2-37). This suggests that these conditions are sufficiently compatible with streptavidin-biotin recognition.

<table>
<thead>
<tr>
<th>whole cell lysate</th>
<th>post-streptavidin enrichment</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>anti-GluA1 (100 kDa)</td>
</tr>
<tr>
<td>2</td>
<td>anti-Homer (41 kDa)</td>
</tr>
<tr>
<td>3</td>
<td>anti-V5 (HRP-NLGN1 is 120 kDa)</td>
</tr>
<tr>
<td>4</td>
<td>same as above, with longer exposure</td>
</tr>
<tr>
<td></td>
<td>anti-PSD95 (80 kDa)</td>
</tr>
</tbody>
</table>

1: Non-denaturing RIPA lysis
2: Denaturing 1% SDS lysis (boiled for 10 min, then diluted into RIPA)
3: Non-denaturing RIPA lysis with additional 4M urea wash during pulldown
4: Denaturing 1% SDS lysis with additional 4M urea wash during pulldown

**Figure. 2-37. Four different lysis/wash conditions were tested following BxXP labeling using HRP-NLGN1.** Blotting of streptavidin-enriched lysates for a cleft marker (GluA1) and intracellular markers (Homer and PSD95) showed that new conditions 2 and 4 effectively removed the latter while retaining the former. Condition 2 was used for large-scale proteomics to map the excitatory and inhibitory synaptic clefts (Chapter 3). Condition 1 was used in previous studies (Hung et al., 2014; Rhee et al., 2013). The anti-V5 blot detects HRP-NLGN1.
We had also tried harsher urea washes (up to 4M) and found that these conditions were not sufficient to break up the complex and eliminate cytoplasmic PSD proteins (compare conditions 1 and 3). Streptavidin itself is known to maintain its tetramer and bind biotin even when exposed to up to 6M urea (Kurzban et al., 1991). In fact with these washes the beads themselves started to aggregate as a result of the urea washing, and we decided that 4M urea might itself be too harsh for the beads themselves.

Targeting the excitatory and inhibitory synaptic clefts

Because the synaptic cleft had never been isolated before and was less characterized than the other sub-compartments of the synapse, we rationalized that by using multiple HRP baits to the synapse and taking the intersection of the proteomes generated by each of them we could improve the specificity of our dataset. Our initial proteome generated using HRP-NLGN1 also revealed the presence of numerous GABAergic inhibitory synapse components (Fig. 2-35), even though endogenous NLGN1 is thought to localize to glutamatergic excitatory synapses. This encouraged us to explore other HRP baits that could differentiate better between excitatory vs inhibitory synapses.

We fused HRP to a number of known synaptic adhesion proteins (Chih et al., 2006; Linhoff et al., 2009; Takahashi et al., 2012; Varoqueaux et al., 2004) that were known to be specific for glutamatergic and GABAergic synapses (the eventual HRP fusions we decided on are shown in Fig. 2-38) and screened the fusions first for their potential to label synapses (as assayed by punctate BxxP labeling) before assaying them for synaptic localization (Fig. 2-39). Because the synapse itself is an open non-membrane bound structure, it is possible for proteins to diffuse into and out of it. We rationalized that proteins that are synaptogenic, while also synaptic adhesion proteins were likely to have a more stable lifetime within the synapse because they are likely to play a structural role in synapses, and would be preferable to synaptic proteins whose localization in the synapse is dynamically regulated (Scannevin and Huganir, 2000) such as for
glutamate receptor subunits. The other criteria we used to determine our choice of synaptic protein fusion was its tolerance for GFP fusions at the N terminus (which does not affect its surface targeting, that in turns affects the surface HRP available for labeling).

![Figure 2-38. HRP fusion constructs developed for the synaptic cleft proteomics experiment.](image)

**HRP constructs specifically target to their respective synaptic cleft subtypes**

We evaluated our fusion construct’s targeting to these synapses by colocalization of proteins biotinylated by HRP with markers of excitatory and inhibitory synapses, shown in green and blue respectively (Fig. 2-39). Colocalization can be observed by yellow spots for the excitatory constructs, and pink spots for the inhibitory constructs, while being absent for the opposing synapse type. For the excitatory cleft, fluorescence imaging showed that both HRP-LRRTM1 and HRP-LRRTM2 co-localized with the excitatory marker vGlut1 but not the inhibitory synapse marker vGAT (Fig. 2-39). The inhibitory fusion constructs HRP-NLGN2A and HRP-SLITRK3 showed the reverse trend (Fig. 2-39). This is quantified in a bar chart (Fig. 2-40), showing LRRTM1 and 2 are specific for excitatory synapses and NLGN2 and SLITRK3 are specific for inhibitory synapses reflecting what is known in the literature about these 4 synaptic proteins.

The HRP fusion to the well-characterized excitatory synaptic adhesion protein neuroligin-1 (Nlgn1) (Song et al., 1999), again did not show clean
targeting to excitatory synapses (Fig. 2-41), and in our previous preliminary proteomics experiment enriched numerous inhibitory synapse proteins (Fig. 2-35). Hence we decided not to use HRP-NLGN1 for the next proteomic experiment.

**Figure. 2-39.** Fluorescence imaging of synaptic HRP fusion constructs with respect to excitatory and inhibitory synapse markers. For maximal detection sensitivity, the HRP constructs are visualized via BxxP labeling for 1 minute followed by staining with neutravidin-AlexaFluor647 (red). Excitatory synapses (green) and inhibitory synapses (blue) are visualized with anti-vGlut1 and anti-vGAT antibodies, respectively. Sites of red-green overlap are colored yellow. Sites of red-blue overlap are colored pink. Scale bars, 10 µm.

**Figure. 2-40.** Quantitation of colocalization extent for images shown in (D) along with 7 other fields of view containing >900 "puncta" per construct. Errors, ±1 s.d.
Figure 2-41. Fluorescence imaging of HRP-NLGN1 (red) with respect to excitatory (anti-vGlut1, green) and inhibitory (anti-vGAT, blue) synapse markers. Experiment was performed and quantified as in Figure 2-39. Scale bar, 10 um.

New HRP fusions are targeted to synapses

In addition to demonstrating that the new HRP fusions were preferentially localized to either excitatory or inhibitory synapses, we also evaluated the fraction of each HRP reporter is localized to the synaptic cleft surface vs non-synaptic sites on the dendritic surface. Colocalization with vGlut1 and vGAT data shown above for the distribution of the synaptic population of reporters to excitatory vs inhibitory sites quantified only the HRP puncta that overlap with either vGlut1 or VGAT; and that a number of puncta might not overlap with neither marker. We believe this is not because they are non-synaptic; it could arise because these puncta are likely at vGlut2 or vGlut3-positive termini that are not stained by the anti-vGlut1 antibody. We then evaluated the synaptic localization of the new HRP constructs by colocalization with Bassoon, a presynaptic protein found within the active zone and found that >85% of HRP reporter puncta overlap with the endogenous pre-synaptic marker Bassoon (Fig. 2-42-43). This shows that the vast majority of each HRP fusion construct is localized to synaptic versus non-synaptic sites.
Figure. 2-42. Fluorescence imaging of synaptic HRP fusion constructs with respect to synapse marker Bassoon. For maximal detection sensitivity, the HRP constructs are visualized via BxxP labeling for 1 minute followed by staining with neutravidin-AlexaFluor647 (green). Synapses (red) are visualized with anti-Bassoon antibody. Sites of red-green overlap are colored yellow. Scale bars, 10 μm.

Figure. 2-43. Quantitation of colocalization extent for images shown in Fig. 2-42. Images in Fig. 2.42 were quantified along with at least 10 other fields of view per construct. Errors, ±1 s.e.m.

We characterized the four HRP fusion constructs shown by electron microscopy as well, in order to assess at nanometer spatial resolution whether HRP was cleanly targeted to the synaptic cleft region. Instead of immunogold staining, we capitalized on HRP's known ability to generate an osmiophilic polymer from diaminobenzidine (Li et al., 2010). Figs. 2-44 and 2-45 show that the dark stain, indicative of HRP activity, is cleanly localized to the cleft region.
within a few nanometers of synaptic vesicles, for each of the four fusion constructs.

Figure. 2-44. Electron microscopic (EM) visualization of HRP fusion constructs. HRP catalyzes the polymerization and local deposition of diaminobenzidine, which in turn recruits electron dense osmium (Martell et al., 2012). The dark-stained regions in the synaptic clefts indicate the presence of HRP. SV, synaptic vesicles. Scale bars, 200 nm. Samples were generated in collaboration with a post-doctoral fellow Dr. Philipp Stawski.

Figure. 2-45. Electron microscopy of synaptic HRP fusion constructs. Same conditions (but different fields of view) as in Fig. 2-44 and zoomed out to show multiple synapses per field of view. Dark stain indicating the presence of HRP is restricted to cleft regions only and not observed in non-synaptic cell surface regions. Images shown are representative of >6 images per
To generate sufficient material for mass spectrometric analysis, we utilized dissociated cultures of \(~10^7\) cortical neurons harvested from E18 embryonic rat pups. Labeling was performed after 19 days of growth in vitro, when fluorescence and electron microscopy showed that mature synapses were abundant. Lentiviral infection gave HRP expression in \(~80\%\) of neurons (Fig. 2-46) but maintained low expression levels that enabled us to observe punctate and synapse-localized BxxP labeling. Streptavidin blot analysis showed that each HRP construct biotinylated numerous endogenous proteins after only a one-minute reaction with BxxP and \(\text{H}_2\text{O}_2\) (Fig. 2-47) and we could enrich successfully for these proteins using the newly developed washing and lysis protocols for dissembling the PSD. Promisingly, the protein banding patterns as visualized by silver staining, resembled the streptavidin HRP blotting pattern, adding confidence to the success of the new enrichment protocol.

![Figure 2-46. Infection efficiency of lentiviruses used for synaptic cleftome experiments.](image)

Neurons were infected at DIV14 under the same conditions using the same viral titers that give synaptic labeling (these are matched conditions for proteomics performed in Chapter 3), then fixed at DIV 19. Cells were permeabilized with triton-X100 and stained for the V5 epitope on each HRP construct (this reveals the total pool of protein), and the neuronal marker NeuN (Mullen et al., 1992).
Since the proteins we had chosen as our HRP baits were known synaptogenic proteins, there were concerns overexpression artifacts such as protein mistargeting outside of synapses could occur (as we had seen when the HRP fusions are expressed using L2K or at high levels) or that grossly perturbed artificial synapses could be formed, we wanted to make sure that the proteins were expressed at reasonable levels and that the neurons had not been perturbed in their synaptic morphology. Since many of the bait proteins we had used as HRP fusions did not have known validated antibodies, we decided to check the overexpression levels by two methodologies, quantitative polymerase chain reaction (qPCR) which measures relative RNA levels and western blotting for the protein. qPCR revealed that at the relative transcript levels (Fig. 2-48), there was only ~1.5 fold more total transcript (including endogenous transcripts which our qPCR probe targets as well as the recombinant) of each of the HRP bait proteins. We were only able to identify one good antibody for detecting the HRP baits at the protein level, for Neuroligin 2, and this showed a less than 0.5 fold overexpression of the protein (Fig. 2-49). Since we knew that at these levels, proteins were expressed at roughly equivalent amounts (shown in Chapter 3 Fig. 3-6) we believe that the overall overexpression levels are not high.
Figure. 2-48. Relative total mRNA levels for each synaptic HRP bait. The relative levels of the endogenous transcript are shown in black and after introduction of HRP-tagged constructs by lentiviral infection shown in grey. qPCR data was generated by Austin Draycott.

Figure. 2-49. Western blotting comparison of the levels of endogenous neuroligin 2 protein compared to exogenously introduced HRP-NLGN2. A band representing the increase in molecular weight after fusion to HRP is visible only in neuronal whole cell lysate that has been infected with HRP-NLGN2 (under the same conditions used for proteomics in Chapter 3), at levels that gave tight synaptic targeting.

We also assayed neurons that had been infected with the HRP constructs for two features, relative synapse density and synapse size, which are known to be perturbed under overexpression conditions of synaptogenic proteins. Neither assay showed any perturbation for either the vGlut1 or vGAT synapses, both for density (number of synapses per unit area in Fig. 2-50) and relative sizes of the synapses themselves (Fig.2-51), suggesting that exogenous addition of our HRP constructs does not affect synapse morphology significantly.
Inhibitory synapse density

Figure. 2-50. Relative synapse density after exogenous introduction of HRP fusions to the synaptic baits. Relative synapse density for vGlut1 (green) and vGAT (red) synapses compared to untransfected neurons. n.s.: no significance.

Figure. 2-51. Relative synapse size after exogenous introduction of HRP fusions to the synaptic baits, Relative synapse sizes for vGlut1 (green) and vGAT (red) synapses compared to untransfected neurons. >600 synaptic puncta were measured. n.s.: no significance.

Conclusions

Since we had addressed the technical challenges that we identified in the first pass with the synaptic cleft proteomes, we decided to proceed with implementing these technical advances in our second pass at the synaptic cleft proteomes, using these to map the proteomes of the excitatory and inhibitory synaptic clefts, described in Chapter 3. From a technological viewpoint, the improvements we have developed here help to advance the peroxidase-based proteomic methodology in several respects. We use non-dividing primary cells, rather than fibroblasts (Rhee et al., 2013), and consequently employ a post-digestion, chemical labeling strategy (iTRAQ) for quantification rather than metabolic labeling via SILAC, which requires protein turnover. We use HRP in
place of APEX2, showing that HRP catalyzes the same labeling chemistry. We aim to use this technology to map an “open” or non-membrane enclosed cellular compartment, while achieving high spatial specificity, utilizing the short “labeling radius” of the peroxidase. Previous demonstrations have been in fully enclosed (mitochondrial matrix (Chen et al., 2015; Rhee et al., 2013)) or partially enclosed (mitochondrial intermembrane space (Hung et al., 2014)) cellular regions. BxxP is introduced as a novel probe that restricts proteomic tagging to the cell surface. The newly developed HRP constructs also allow us to differentiate between the excitatory and inhibitory synaptic clefts. Using the new streptavidin enrichment protocol we can now disassemble the post-synaptic density and allows the clean capture of proteins directly biotinylated by HRP.

Materials and Methods

Characterization of peroxidase fusion constructs

Localization and activity of peroxidase fusions by fluorescence microscopy

Cortical neurons were cultured as described under “Rat cortical neuron culture”. Cortical neurons at days in vitro (DIV) 15 were infected with lentiviruses expressing the HRP-tagged synaptic constructs at a titer previously empirically optimized (see “Titrating lentiviral constructs”) to give tight, localized BxxP labeling.

Four days later, at DIV 19, cells were treated with 100 uM BxxP and 1 mM H2O2 in Tyrode’s Buffer (145 mM NaCl, 1.25 mM CaCl2, 3 mM KCl, 1.25 mM MgCl2, 0.5 mM NaH2PO4, 10 mM glucose, 10 mM HEPES, pH 7.4) for 1 min at room temperature, and then quenched with 3 washes of Tyrode’s buffer containing 10 mM sodium azide, 10 mM sodium ascorbate, and 5 mM Trolox. The samples were then fixed with 4% paraformaldehyde in “fixation buffer” (60 mM PIPES, 25 mM HEPES, 10 mM EGTA, 2 mM MgCl2, 0.12 M sucrose, pH 7.4) at room
temperature for 10 minutes. The cells were washed three times with Dulbecco’s Phosphate Buffered Saline (DPBS).

For demonstrating the impermeability of BxxP, a slightly modified labeling protocol was used. Cortical neurons at days in vitro (DIV) 15 were infected with lentiviruses expressing the HRP-tagged synaptic constructs or APEX2 tagged constructs. At DIV19, 100 μM BxxP was preincubated with cells in Tyrode’s buffer for 30 min at 37 °C. Then 1 mM H₂O₂ was added for 1 minute at room temperature. Samples were quenched and fixed as above. For BP labeling, cells were incubated with 500 μM Biotin-phenol (BP) in Tyrode’s buffers for half an hour, or added together with 1 mM H₂O₂ for 1 min. Samples were then quenched and fixed as above.

To detect extracellular biotinylated proteins, cells were blocked at room temperature with 3% w/v bovine serum albumin (BSA) in DPBS for one hour at room temperature or overnight at 4 °C, then stained with neutravidin protein (Invitrogen) pre-coupled to AF647-NHS esters (Invitrogen), at a 1:1000 dilution for one hour, then washed three times with DPBS. Cells were post-fixed for 10 min with fixed with 4% paraformaldehyde in “fixation buffer” (60 mM PIPES, 25 mM HEPES, 10 mM EGTA, 2 mM MgCl₂, 0.12 M sucrose, pH 7.4) at room temperature, then permeabilized as described below for staining endogenous antibodies.

Cells were permeabilized with 0.1% TritonX-100 in DPBS for 7 minutes at room temperature. After three more washes with DPBS, the cells were blocked with 3% w/v bovine serum albumin (BSA) in DPBS for one hour, then stained with mouse anti-V5 (Invitrogen, 1:1000 dilution), or mouse anti-vGlut1 (NeuroMab, 1:1000) and rabbit anti-vGAT (Synaptic Systems, 1:1000), mouse anti-Bassoon (Enzo Life sciences 1:1000), in 3% w/v BSA in DPBS for 2 hour at room temperature. Samples were then washed 3 x 1 minutes with DPBS. Samples were incubated with secondary goat anti-mouse-Alexa-Fluor-488 (AF488) and
goat anti-rabbit-Alexa-Fluor-568 antibody (AF568) (Invitrogen, 1:1000 dilution each) in 3% w/v BSA in DPBS for one hour, then washed 3 x 1 minutes with DPBS. Samples were imaged by confocal microscopy with 5 x 0.3 um stacks using a 100x objective. Images shown in figures are z-projections of all 5 stacks acquired per field of view.

**Generation of biotinylated cell lysate for SDS-PAGE analysis**

Two 10 cm dishes were prepared with five million cortical neurons in each. At DIV 15, neurons were infected with lentiviruses expressing the HRP-tagged synaptic constructs, using a previously optimized titer (see “Titration of lentivirus ”). At DIV 19, 100 µM BxxP and 1 mM H₂O₂ in Tyrode’s Buffer were added for 1 min at room temperature. The reaction was quenched with 3 washes of Tyrode’s buffer containing 10 mM sodium azide, 10 mM sodium ascorbate, and 5 mM Trolox. After quenching, the cells were harvested by scraping and pelleted by centrifugation at 3,000 g for 10 minutes. The supernatant was discarded and the pellet was stored at -80 °C overnight.

The cell pellet was lysed by incubating in 100 µl of 1% SDS lysis buffer (50 mM Tris-HCl, 1% SDS, pH 8.0) containing protease inhibitor cocktail (Sigma Aldrich, catalog no. P8849), 1 mM PMSF (phenylmethylsulfonyl fluoride), 10 mM sodium azide, 10 mM sodium ascorbate, and 5 mM Trolox, then boiling for 5 min to denature and separate postsynaptic density proteins. The sample was then diluted in 400 µL of 1.25X RIPA lysis buffer to give 1X RIPA lysis buffer (50 mM Tris-HCl, pH 8.0, 150 mM NaCl, 0.2% SDS, 0.5% sodium deoxycholate, 1% Triton-X 100). Lysates were centrifuged at 16,000 g for 10 minutes at 4 °C, with the supernatant used for subsequent processing.

**Streptavidin blotting of biotinylated neuron lysate**
Lysates were combined with SDS protein loading buffer and boiled for 5 min, then run on a 8% SDS-PAGE gel. The gel was transferred to a nitrocellulose membrane. The membrane was stained with Ponceau S (10 minutes in 0.1% w/v Ponceau S in 5% (v/v) acetic acid/water) and imaged. After destaining in deionized water, the membrane was blocked with 3% w/v BSA in 1 x TBST (0.1% Tween-20 in Tris-buffered saline) at 4 °C overnight. The membrane was rocked in 0.3 μg/ml streptavidin-HRP (Thermo Scientific) in 1% w/v BSA in 1 x TBST at room temperature for 1 hour, then washed with 1 x TBST four times for 5 minutes each time. Finally, the blot was developed with Clarity Western ECL Substrate (Bio-Rad) and imaged on an Alpha Innotech gel imaging system.

Streptavidin enrichment of biotinylated lysate and analysis by silver staining

To enrich biotinylated proteins, streptavidin-coated magnetic beads (Pierce catalog no. 88817) were prepared by washing twice with RIPA lysis buffer. Then, 500 μl of whole cell lysate was incubated with 150 μl of streptavidin bead slurry overnight at 4 °C with gentle rotation. The beads were washed with 2 × 1 ml RIPA lysis buffer, 1 x 1 ml of 1M KCl, 1x 1 ml of 0.1 M Na2CO3, 1 x 1 ml of 2 M urea in 10 mM Tris-HCl, pH 8.0, and again with 2 × 1 ml RIPA lysis buffer.

Biotinylated proteins were eluted by boiling the beads for 10 min in 50 μl 3x protein loading buffer supplemented with 20 mM dithiothreitol (DTT) and 2 mM biotin. The streptavidin eluate (SAE) was collected and run on a 8% SDS-PAGE gel. To analyze biotinylated proteins, western blots were performed as described above. To analyze all protein material eluted, gels were silver stained (Pierce) instead, then imaged on an Alpha Innotech gel imaging system.

Optimization of neuron lysis conditions
To identify lysis and pulldown conditions that separate cytoplasmic proteins from biotinylated cleft-exposed synaptic proteins a number of variations in the lysis and pulldown washes were tried.

Two 10 cm dishes of containing 5 million DIV19 cortical neurons infected with HRP-NLGN1 at DIV 15 were first labeled as described in "HRP mediated biotinylation to generate whole cell lysates for streptavidin-HRP western blotting or streptavidin bead enrichment". Cell pellets were divided equally into four and were lysed either by incubation in:

A) 1X RIPA lysis buffer (50 mM Tris-HCI, pH 8.0, 150 mM NaCl, 0.2% SDS, 0.5% sodium deoxycholate, 1% Triton-X 100) containing protease inhibitor cocktail (Sigma Aldrich, catalog no. P8849), 1 mM PMSF (phenylmethylsulfonyl fluoride), 10 mM sodium azide, 10 mM sodium ascorbate, and 5 mM Trolox for 30 min on ice, or
B) Incubating in 100ul of 1% SDS lysis buffer, (50 mM Tris, pH = 8) containing protease inhibitor cocktail (Sigma Aldrich, catalog no. P8849), 1 mM PMSF (phenylmethylsulfonyl fluoride), 10 mM sodium azide, 10 mM sodium ascorbate, and 5 mM Trolox, then boiled for 10 min. This was then diluted in 400ul of 1.25X RIPA lysis buffer to give 1X RIPA lysis buffer (50 mM Tris-HCI, pH 8.0, 150 mM NaCl, 0.2% SDS, 0.5% sodium deoxycholate, 1% Triton-X 100).

Lysates from A or B were centrifuged at 16,000 g for 10 minutes at 4 °C, then 150 ul of streptavidin bead slurry which had been washed twice with RIPA lysis buffer, was added and incubated overnight at 4 °C with gentle rotation. Beads were then washed with C) 2 × 1 ml RIPA lysis buffer, 1 × 1 ml of 1M KCl, 1x 1 ml of 0.1 M Na2CO3, 1 × 1 ml of 2 M urea in 10 mM Tris-HCl, pH 8.0, and again with 2 × 1 ml RIPA lysis buffer, or D) with an additional 4 M urea in 10 mM Tris-HCl, pH 8.0 wash following the 2 M urea wash. The following combination of conditions were performed for the following lanes: 1) A + C, 2) B + C, 3) A + D, 4) B + D.
Beads were then eluted by boiling the beads for 10 min in 50 µl 3x protein loading buffer supplemented with 20 mM dithiothreitol (DTT) and 2 mM biotin. The streptavidin eluate (SAE) was collected and run on a 8% SDS-PAGE gel.

To analyze biotinylated proteins, western blots were performed as described above; eluates mixed with 1 x SDS-loading buffer and run on a 8% SDS-PAGE gel. The gel was transferred to a nitrocellulose membrane. The membrane was stained with Ponceau S (10 minutes in 0.1% w/v Ponceau S in 5% acetic acid/water) and imaged. After destaining in deionized water, the membrane was blocked with 3% w/v BSA in 1 x TBST (0.1% Tween-20 in Tris-buffered saline) at 4 °C overnight. The membrane was rocked with one of the following primary antibodies: mouse anti-GluA1 (NeuromAb, 1:1000 dilution), rabbit anti-Homer (SynapticSystems, 1:1000 dilution), mouse anti-V5 (Invitrogen, 1:1000 dilution) or mouse anti-PSD95 (Neuromab, 1:1000 dilution) in 1% w/v BSA in 1 x TBST at room temperature for 1 hour, then washed with 1 x TBST four times for 5 minutes each time. The membrane was then rocked with Goat anti-Mouse HRP conjugate (1:3000 dilution) or Goat anti-rabbit HRP conjugate (1:3000 dilution) depending on the primary antibody used, for 1 hour, then washed with 1 x TBST four times for 5 minutes each time. Finally, the blot was developed with Clarity Western ECL Substrate (Bio-Rad) and imaged on an Alpha Innotech gel imaging system.

**Neuron cultures and lentiviral infection**

**Dissociated rat cortical neuron culture**

Sprague Dawley rat embryos were sacrificed at embryonic day 18. Dissected cortical tissue was digested with papain (Worthington) and DNasel (Roche), then plated on 0.09–0.12 mm thickness glass coverslips (Carolina Biological Supply) in a 1:1 volume ratio of growth medium A and growth medium B and cultured at 37°C under 5% CO₂. Growth medium A is MEM (Sigma) with L-glutamine (Sigma) supplemented with 10% (v/v) fetal bovine serum (PAA laboratories) and
2% (v/v) B27 (Life Technologies). Growth medium B is Neurobasal medium (Life Technologies) supplemented with 2% (v/v) B27 and 1% (v/v) GlutaMAX (Life Technologies). Glass coverslips were pretreated with poly-D-lysine (Sigma) and mouse laminin (Life Technologies). At 4 days in vitro, half of the spent culture medium was replaced with fresh growth medium B in addition to 10 μM FUDR (Sigma), an anti-mitotic, and replaced similarly every 4 days onwards.

Preparation of lentivirus

To prepare lentiviruses, human embryonic kidney (HEK) 293T cells were plated in a T25-flask (or scaled appropriately), and cultured in Minimum Essential Medium (MEM) supplemented with 10% fetal bovine serum, penicillin, and streptomycin at 37°C under 5% CO₂. Each flask of cells was transfected with 2.25 μg of FSW plasmid, 0.25 μg VSVG, and 2.0 μg dR8.91 using 20 μl polyethelenimine (Sigma) in MEM (without serum or antibiotics) at ~70% confluence. VSVG and dR8.91 are lentiviral packaging plasmids (Pagliarini et al., 2008). The cells were transfected for 2-3 hours, then media was replaced with 5 ml fresh growth media. After 48 hours, the supernatant was collected and filtered through a 0.45 μm syringe filter, and flash-frozen with liquid nitrogen and stored at -80 °C.

To infect cells, frozen virus was thawed quickly in a 37°C water bath, and added to cells at an MOI that gave synaptic punctate localization (previously measured in a separate titration experiment). Cells were infected for four days to allow robust expression and trafficking.

Titration of lentivirus

Lentiviruses were first generated as described under "Generation of lentiviruses for trans-gene expression in cortical neurons and infection protocol". Neurons were infected at DIV 15 with various titers of lentiviral supernatant typically
ranging from 200 ul to 1 ul). At DIV 19, cells were labeled with 100 μM BxxP and 1mM H₂O₂ in Tyrode’s Buffer for 1 min at room temperature, and then quenched with 3 washes of Tyrode’s buffer containing 10 mM sodium azide, 10 mM sodium ascorbate, and 5 mM Trolox. The samples were then fixed with 4% paraformaldehyde in “fixation buffer” at room temperature for 10 minutes. The cells were then washed three times with Dulbecco’s Phosphate Buffered Saline (DPBS).

Expression levels and multiplicity of infection were inferred from BxxP labeling. The sample was blocked at room temperature with 3% w/v BSA in DPBS for one hour at room temperature or overnight at 4 °C, then stained with neutravidin protein (Invitrogen) pre-coupled to AF647-NHS esters (Invitrogen), at a 1:1000 dilution for one hour. Samples were then washed three times with Dulbecco’s Phosphate Buffered Saline (DPBS). Samples were then imaged by confocal microscopy using a 40x objective. The optimal titer for proteomics and imaging was determined by selecting a titer in which BxxP labeling was not diffuse and mostly punctate and more than 80% of cells in the dish were infected (as determined by comparing the BxxP labeling detected by fluorescence imaging to the DIC images). Note that only synaptic targeted constructs gave a punctate tight labeling pattern, while general surface targeted constructs gave a diffuse localization, for example HRP-TM.

Synthesis and characterization of BxxP

Synthesis of BxxP

All starting materials were purchased from Sigma Aldrich. BxxP was synthesized as follows: Biotinamidohexanoyl-6 amino hexanoic acid N-hydroxysuccinimde ester (150 mg) and 36.2 mg (1.0 equivalents) of tyramine were dissolved in 2.7 ml of dimethylsulphoxide (DMSO). 276 ul or 6.0 equivalents of DIPEA (N,N-diisopropylethylamine) was then added, and the reaction was incubated
overnight at room temperature with stirring. To work up the reaction, 15 ml of
H\textsubscript{2}O was added and the reaction was then frozen and lyophilized overnight to
remove DMSO. A solid mixture of a brown and white precipitate remained. 7 ml
of cold methanol at -20 °C was then added drop-wise until all the brown
precipitant had dissolved, and the reaction was chilled at -20 °C for another 3
hours, leaving only a white precipitate. The precipitate was quickly filtered over a
frit funnel, and washed with ethyl acetate 4 times. The precipitate was then dried
in vacuo yielding 117 mg of BxxP.

NMR Characterization of BxxP

\textsuperscript{1}H NMR spectra was obtained on a Bruker Advance 400 MHz spectrometer and
processed on TopSpin 1.3 software.

\textsuperscript{1}H-NMR for BxxP (400 MHz, DMSO-d\textsubscript{6}): \textit{d} = 9.17 (s, 1H), 7.89-7.62 (m, 3H), 6.97
(d, \textit{J}\textsubscript{3} = 8.4 Hz, 2H), 6.66 (d, \textit{J}\textsubscript{3} = 8.4 Hz, 2H), 6.43 (s, 1H), 6.37 (s, 1H), 4.36-
4.26 (m, 1H), 4.19-4.08 (m, 1H), 3.23-2.89 (m, 6H), 2.81 (dd, \textit{J}\textsubscript{2} = 12.2 Hz, \textit{J}\textsubscript{3} =
5.2 Hz, 1H), 2.75-2.68 (m, 1H), 2.61-2.53 (m, 3H), 2.08-1.96 (m, 6H), 1.66-1.03
(m, 18H) ppm.

Mass Spectrometry characterization of BxxP

MS characterization of purified BxxP on an Agilent 6500 series Q-TOF LC/MS
instrument showed: BxxP, calculated for C\textsubscript{30}H\textsubscript{47}N\textsubscript{5}O\textsubscript{5}S (M+H): 590.33; found:
590.327.

Fluorescence microscopy

Fluorescence microscopy
Fluorescence confocal microscopy was performed with a Zeiss AxioObserver inverted microscope with 40× and 100× oil-immersion objectives, outfitted with a Yokogawa spinning disk confocal head, Cascade II:512 camera, a Quad-band notch dichroic mirror (405/488/568/647), and 405 (diode), 491 (DPSS), 561 (DPSS) and 640 nm (diode) lasers (all 50 mW). CFP (405 laser excitation, 445/40 emission), Venus/Alexa Fluor 488 (491 laser excitation, 528/38 emission), Alexa Fluor 568 (561 laser excitation, 617/73 emission), and AlexaFluor 647 (640 laser excitation, 700/75 emission) and differential interference contrast (DIC) images were acquired through a 100× or 40x oil-immersion lens using Slidebook (Intelligent Imaging Innovations). All image analysis was performed in Slidebook. Fluorophore channels in each experiment were normalized to the same intensity ranges. Acquisition times ranged from 300ms to 1000ms. Neuron images were projection summations from 0.3 μm-step optical stacks spanning 1.5 μm total depth.

Image analysis

For quantification, Slidebook 5.0 software (Intelligent Imaging Innovations) was first used to calculate the mean intensities and the standard deviations for each channel for every field of view (FOV). To generate masks for each channel with consistency and in an unbiased manner, 2x of the standard deviation was added to the mean intensity for that FOV, and this value was used to determine the threshold used to generate a mask for "real signal" for that channel in that FOV.

To analyze the synaptic density a mask for the venus channel was crossed with either a mask for the vGlut1 or vGAT channel to generate new masks of their overlapping regions. The fluorescence intensity of the vGlut1 or vGAT channels that overlapped with the venus channel was averaged over the area of the entire venus mask, giving the average intensity of the synaptic marker recruited over the transfected neuron. This was then normalized to the control condition to obtain the relative densities.
To analyze the overlap of BxxP labeling with synaptic markers a mask for the BxxP detected by neutravidin-647 channel was crossed with either a mask for the vGlut1 or vGAT channel to generate new masks of their overlapping regions. This mask allowed us to count the number of “puncta” that overlapped with either the vGlut1 or vGAT channel. The % colocalization with synaptic marker reported is the number of puncta that overlap with the marker divided by the sum of all of the vGlut1 and vGAT overlapping puncta.

**Electron microscopy**

Cortical neurons were cultured as described under “Rat cortical neuron culture”, however cells were plated in laminin and poly-D-lysine coated 35mm glass bottom dishes (MatTek).

Cortical neurons at DIV 15 were infected with lentiviruses expressing the HRP-tagged synaptic constructs at an MOI previously determined (see “Titrating lentiviral constructs”) to give tight, localized synaptic BxxP labeling of synapses in a separate experiment. 7uM of heme was added to the culture media at DIV 18 to improve peroxidase activity.

At DIV 19, fixation and staining for EM was performed similarly as previously described (Martell et al., 2012; Rhee et al., 2013). For fixation, 37 °C 2% glutaraldehyde (Electron Microscopy Sciences) in sodium cacodylate buffer (100 mM sodium cacodylate with 2 mM CaCl2, pH 7.4) was added to each sample, this was removed immediately and fresh fixative was added again and then the cells were moved to ice for 60 minutes. The samples remained on ice through the uranyl acetate staining. After fixation, the samples were washed 5 × 2 minutes in cold sodium cacodylate buffer, quenched with 20 mM glycine in sodium cacodylate buffer for 5 minutes, and rinsed again with cold sodium cacodylate buffer for 5 × 2 minutes. The samples were then reacted with a
solution of 1.4 mM 3,3'-diaminobenzidin (DAB) and 1 mM H$_2$O$_2$ in cold sodium cacodylate buffer to allow HRP to catalyze the polymerization of DAB for 30 minutes up to an hour depending on the construct. The cells were then rinsed 5 × 2 minutes with cold sodium cacodylate buffer.

The DAB polymers were subsequently stained with reduced 1% OsO$_4$ (R-OTO) for 2 hours in cold sodium cacodylate buffer. R-OTO was made by mixing 2% potassium ferrocyanide (Alfa Aesar) in sodium cacodylate buffer with 2% OsO$_4$ (Electron Microscopy Sciences) and chilling to 4 °C. Samples were rinsed 5 × 2 minutes in chilled Millipore water. Cold 2% aqueous uranyl acetate (Electron Microscopy Sciences) was then added, and the samples were incubated overnight at 4 °C.

The samples were rinsed 5 x 2 min with distilled water, then dehydrated for 2 minutes each in 20%, 50% 70%, 90%, 95%, 100% at 4 °C, and 100% ethanol at room temperature.

Samples were infiltrated for half an hour in Durcupan ACM resin (Sigma Aldrich) using 1:1 (v/v) resin and anhydrous ethanol for 30 minutes, then 4 exchanges of 100% resin for 1-2 hours each. After infiltration, the samples were polymerized for 48 hours at 60 °C. The DAB-stained areas of the embedded cell samples were identified by transmitted light, selectively sawed out using a jeweler's saw and mounted on dummy acrylic blocks with cyanoacrylic adhesive (Krazy Glue, Elmer's Products). The coverslip was removed, the block trimmed, and cut into ultrathin 80 nm sections using a diamond knife on a Leica Ultracut UCT. Sections were mounted on graphene-covered copper grids and imaged at the Whitehead Institute Keck Microscopy Facility on a FEI Tecnai G$^2$ Spirit BioTWIN transmission electron microscope operated at 80 kV or using a JEOL 1200 TEM operating at 80 keV.

Other chemicals and reagents
BP labeling and reagents are described in reference (Hung et al., 2016; Rhee et al., 2013). All chemicals were purchased from Sigma-Aldrich unless otherwise specified.

References


Green, N.M. (1975). Advances in Protein Chemistry Volume 29 (Elsevier).


Chapter 3: The excitatory and inhibitory synaptic cleft proteomes

Sections of this chapter are excerpts taken from a submitted manuscript: K. H. Loh, P. S. Stawski, N. D. Udeshi, T. Svinkina, T. Deerinck, M. H. Ellisman, S. A. Carr, and A. Y. Ting. Proteomic maps of the excitatory and inhibitory synaptic clefts reveal specificity factor for pre-to-post synaptic matching. Electron microscopy was kindly performed in collaboration with Philipp Stawski and Tom Deerinck in Mark Ellisman’s lab at UCSD. Proteomics was performed in collaboration with Namrata Udeshi and Tanya Svinkina in Steven Carr’s lab at the Broad Institute.
Introduction

In contrast to the previous sub-mitochondrial compartments we have mapped (Hung et al., 2014; Rhee et al., 2013), the synaptic cleft is an open, non-membrane enclosed structure that pushes the spatial resolution limits of our peroxidase technology. Our goal was to confidently distinguish between bona fide cleft-resident endogenous proteins and extra-synaptic cell surface proteins that might be only nanometers away from the synapse. In conjunction with insights gained from the preliminary synaptic cleft proteome using HRP-NLGN1 (see Chapter 2) and the new constructs and methodologies developed to overcome the challenges encountered there, we proceeded to map the proteomes of the excitatory and inhibitory synaptic clefts.

Proteomic mapping of the excitatory and inhibitory synaptic cleft

*iTRAQ experimental design for mapping the excitatory and inhibitory synaptic clefts*

With the new HRP fusions that we had developed and characterized, 2 that would target to the glutamatergic excitatory synapse and 2 to the GABAergic inhibitory synapse, we designed and carried out a 4 state quantitative proteomics by iTRAQ using the following experimental setup (Fig. 3-1). Each experiment has one excitatory cleft targeted construct and an inhibitory cleft targeted one, which could help us distinguish between the two molecularly distinct synapse types. The general cell surface targeted construct “HRP-TM” would be used to differentiate between proteins generally on the cell surface versus actually being within synaptic clefts. Lastly a negative labeling control would be used to identify only genuinely biotinylated proteins and eliminate non-specific bead binders.

The workflow for the proteomics experiment is similar to that applied to the proteomic mapping when we used HRP-NLGN1 as bait: rat cortical neurons were cultured until DIV 15, infected with the respective lentiviruses, and cultured for
another 4 days before proteins are biotinylated at DIV 19 (a detailed labeling protocol is provided at the end of this chapter). Proteins are labeled with BxxP and H$_2$O$_2$, followed by quenching of the labeling. Cell are scraped and pelleted, and the neuronal cell pellets are lysed using the newly developed protocol for dissociating the PSD (Fig. 2-36 in Chapter 2). Protein inputs are normalized using a BCA assay before streptavidin magnetic beads are added for enrichment, and 7mg of protein (6 10cm dishes, 15 million cortical neurons each) was used for each state before adding streptavidin beads (450ul) for enrichment overnight at 4 °C. Beads were then washed similarly to mapping the synaptic cleft proteome with HRP-NLGN1 and provided to our mass spectrometry collaborators (Dr. Namrata Udeshi and colleagues at the Broad Institute) who digested proteins with trypsin on beads. The eluted peptides were tagged with unique iTRAQ labels, and liquid chromatography and tandem mass spectrometry (MS) was then performed on a mixture of the four tagged states (a complete protocol is provided in the methods section).

**Figure 3-1.** Ratiometric tagging strategy for mapping the excitatory and inhibitory synaptic clefts using iTRAQ. Each experiment consisted of four samples; HRP targeted to either the excitatory or the inhibitory synaptic cleft, a general cell surface targeted HRP and a negative labeling control. Comparisons between states would allow identification of non-specific bead binders, synaptic cleft proteins residing within each type of cleft and general cell surface proteins.
In the mass spectrometer, each iTRAQ reagent is fragmented and generates reporter groups that report on the relative quantities between states at the MS/MS level (Fig. 3-2). Simply, a non-specific bead binder would have the following fragmentation profile; it is not enriched to a greater extent by the HRP bait than the negative control. A general cell surface protein that is not synaptic would be enriched equally in the 3 states which HRP is targeted on the cell surface (either synaptic or the general cell surface), while an excitatory cleft protein would be enriched in the red state (based on the colors shown in Fig. 3-1 where HRP is targeted to the excitatory synapse), but not in the blue (inhibitory state), yellow (general cell surface) and even less so than the green (negative control). Analogously, an inhibitory synaptic cleft protein would be enriched by the blue state instead of the red one. It is also plausible that there are proteins that are found in both synapse types (for eg. such as the neurexins) and these might be enriched equally by both the red and blue states, but more so than the yellow and the green.

![Figure 3-2. Expected iTRAQ reporters fragment profiles at the MS² level for various types of proteins in our proteomic data. Cartoon representation of iTRAQ reporter ion fragmentation profile expected from non-specific bead binders, general surface proteins and excitatory synaptic cleft proteins. Colors represent state iTRAQ reporter is derived from (see Fig. 3-1).](image)

Since we had 2 HRP fusions per synapse subtype, we ran a total of 3 proteomic experiments with different permutations of these HRP constructs (Fig. 3-3). Our expectation is that these permutations would improve the specificity of the proteomes generated and average out any natural bias with regards to labeling endogenous binding partners of the HRP baits that might not be synaptic (such a situation could arise if the protein forms a dimeric complex with other proteins that might have a different synaptic preference, for eg. neuroligin 1 is known to form heterodimers with other neuroligin members in vitro (Poulopoulos...
et al., 2012), even each neuroligin isoform has a unique set of synapse types that it localizes to (ie Nlgn1 to excitatory synapses, Nlgn2 to inhibitory synapses).

<table>
<thead>
<tr>
<th>Sample</th>
<th>Experiment 1</th>
<th>Experiment 2</th>
<th>Experiment 3</th>
<th>iTraq label</th>
</tr>
</thead>
<tbody>
<tr>
<td>Biotinylation by excitatory cleft-targeted HRP</td>
<td>HRP-LRRTM1</td>
<td>HRP-LRRTM2</td>
<td>HRP-LRRTM1</td>
<td>114</td>
</tr>
<tr>
<td>Biotinylation by inhibitory cleft-targeted HRP</td>
<td>HRP-NLGN2A</td>
<td>HRP-NLGN2A</td>
<td>HRP-SLITRK3</td>
<td>115</td>
</tr>
<tr>
<td>Biotinylation by general cell surface HRP-TM</td>
<td>HRP-TM</td>
<td>HRP-TM</td>
<td>HRP-TM</td>
<td>116</td>
</tr>
<tr>
<td>No biotinylation</td>
<td>Omit HRP</td>
<td>Omit BxxP</td>
<td>Omit H2O2</td>
<td>117</td>
</tr>
</tbody>
</table>

**Figure 3-3.** Design of three independent proteomics experiments. Each experiment consisted of four samples, which were separately enriched and tagged with unique iTraq labels (right). Mass spectrometry (MS) was performed on the mixture of four samples, resulting in four mass-shifted peaks of varying intensity per detected peptide.

**Fluorescence microscopy imaging as matched experimental controls for proteomics experiment**

We also performed matching fluorescence microscopy imaging as a matched control for titer, BxxP labeling and expression for each of the HRP baits and the negative controls (Fig. 3-4). Because the surface pool of HRP-tagged constructs is undetectable at these expression levels (a sensitivity comparison of HRP vs other epitope tags is shown in Fig. 3-15), we first carried out BxxP labeling, quenched, and fixed cells with 4% paraformaldehyde. Cells were then blocked with 3% BSA and stained with neutravidin-647 conjugate, followed by a second fixation. Cells were washed again and then permeabilized with 0.1% triton-X 100, and the cells were stained for the V5 epitope tag which would reveal the total pool of the HRP baits we had introduced by lentiviruses (which is detectable), marking the neurons that had been successfully infected. Note that at these lentiviral titers for infection and expression, the biotinylation pattern is mostly punctate (and suggestive of the synaptic labeling we had characterized earlier, while labeling for HRP-TM was mostly diffuse and covered the entire cell surface, demonstrating a difference in the pools of proteins we were labeling with these constructs.
Figure 3-4. Imaging the biotinylation activities of synaptic HRP fusion constructs. Fluorescence imaging of the biotinylation catalyzed by HRP fusion constructs used for proteomics. DIV19 rat cortical neurons expressing the constructs indicated across the top were treated with BxxP and H₂O₂ for 1 minute, then fixed and stained with neutravidin-AlexaFluor647 to detect biotinylated proteins (middle row) and anti-V5 antibody (AlexaFluor 488 readout) to detect HRP construct expression (top row). The last two columns show negative controls with H₂O₂ or BxxP omitted. Images below show zoomed in regions of BxxP labeling (dotted lines), demonstrating the difference between punctate (synaptic constructs) and diffuse (HRP-TM) staining patterns. Scale bars, 10 μm. DIC, differential interference contrast.

Western blotting of whole cell lysates of proteins biotinylated by the different HRP constructs as matched experimental controls for the proteomics experiment

We ran streptavidin-HRP western blots of the whole cell lysates corresponding to each experimental state (Fig. 3-5) These gave unique biotinylation patterns between excitatory, inhibitory and the TM constructs, and only the endogenous biotinylated proteins could be observed in the negative labeling control (Chapman-Smith and Cronan, 1999). Of note however is that the intensity of biotinylation between states seemed different, and this could reflect a larger number of endogenous proteins present at excitatory synapses compared to inhibitory synapses, or differential levels of labeling of the HRP baits, although images presented in the fluorescence imaging control (Fig. 3-4) were normalized
and did not suggest this was the underlying factor. We ran a western blot to compare the expression levels of the various HRP baits, and found that there was no significant difference in the expression levels between baits (Fig. 3-6), suggesting that the former might be the explanation for this observation.

Figure 3-5. Analysis of biotinylated cell lysates for proteomics experiments. Streptavidin blot analysis of whole cell lysates of neurons infected and labeled with BxxP. The three bands present in the negative control lanes at 129, 81, and 79 kDa are endogenously biotinylated proteins (Chapman-Smith and Cronan, 1999).

Figure 3-6. Expression levels of HRP baits used for proteomics experiments. Mo anti-V5 western blot analysis of whole cell lysates of neurons infected with each HRP construct for proteomics. Beta actin is stained as a loading control.
Determining cutoffs the final synaptic cleft proteomes

Characteristics of the proteomic data

It was promising to see that in each of the 3 proteomics experiments more than 2400 unique proteins had 2 or more unique peptides identified (a standard used in proteomics for successful detection of a protein), compared to the 770 successful detections in the previous proteomic attempt with HRP-NLGN1. This was more than 1000 proteins higher than the first experiment with HRP-NLGN1 and although not indicative of whether the experiment had succeeded in identifying the proteins within the synaptic cleft, suggested that at least the initial coverage of the proteome would be higher.

![Histogram](image)

**Figure 3-7. Histogram of proteins identified in the proteomics experiment based on iTRAQ 114/117 (left) 115/117 (right) ratio.** A bimodal distribution for total proteins (in black) detected in the proteomic experiment is observed when plotted against the ratio measuring enrichment by HRP targeted to the synaptic cleft compared to the negative labeling control. Histogram in red represents the distribution of intracellular proteins with respect to the same iTRAQ ratio. Data for experiment 1 is shown here.

Because the biotin-phenoxyl radical does not cross membranes, we had found that in the previous APEX derived proteomes (Hung et al., 2014; Rhee et al., 2013), the distribution of proteins with respect to their enrichment ratios for extent of biotinylation would generate a bimodal distribution of proteins if the structure where labeling is occurring is completely membrane bound (such as the mitochondrial matrix). When the structure is semi-porous (such as the IMS), a continuous distribution is obtained instead. Since the synaptic cleft is extracellular and separated from cytoplasmic proteins by the plasma membrane, the expectation is that it should more closely replicate the bimodal distribution
obtained for the mitochondrial matrix. When we plotted the distribution of proteins based on extent of biotinylation (114/117 ratio or the 115/117 ratio), we observed this expected bimodal distribution (Fig. 3-7). Based on this observation, the right leaning peak could potentially comprise surface proteins truly biotinylated by HRP. We also plotted this distribution in comparison to proteins identified as being intracellular (details of how this list was compiled is described below) and the distribution of the enrichment ratios of these proteins corresponded to the left leaning peak.

We next plotted a similar histogram, this time with the enrichment ratios of the true positive list, which comprised a list of proteins that were found in synapses (a detailed explanation of the derivation of this list is described shortly), in comparison to the true negative list. This showed a clear enrichment of these synaptic proteins versus the intracellular proteins. This result however could be similarly obtained however from a pure enrichment of just surface proteins, since all proteins in the true positive list are in fact, surface proteins as well, therefore further analysis was necessary to determine if the experiment had indeed succeeded. Since the data remained promising, we focused on analyzing the data to determine the synaptic cleft proteomes.

![Figure 3-8. Histogram of synaptic cleft exposed proteins and intracellular proteins in the proteomics experiment based on iTRAQ 114/117 (left) 115/117 (right) ratio. Distribution for true positives (synaptic cleft proteins in green) plotted together with a separate distribution for true negatives (intracellular proteins in red), against the ratio measuring enrichment by HRP targeted to the synaptic cleft compared to the negative labeling control, shows that true positive proteins are enriched relative to the true negatives and well separated. Data for experiment 1 is shown here.](image)

Filter analysis for synaptic cleftomes
To determine the synaptic cleft proteomes, we employed a series of filtering steps (discussed individually in the following sections) that uses the quantitation provided by iTRAQ to identify a final high confidence dataset that would comprise the synaptic cleft proteomes. According to the scheme in Fig. 3-1, the 114/117 and 115/117 iTRAQ ratios for each MS-detected peptide reflect the extent of biotinylation by HRP and can be used to filter out non-specific bead binders. The 114/116 iTRAQ ratio reflects the extent to which a protein is preferentially biotinylated by HRP in the excitatory cleft versus general cell surface HRP-TM. The 115/116 iTRAQ ratio for inhibitory synapse enrichment is analogous. We found that the use of ratiometric iTRAQ filtering, with HRP-TM as the background or reference construct, greatly improved the spatial specificity of our dataset, similarly to the IMS proteome (Hung et al., 2014). Enriched protein lists from the three independent replicates (which utilized different excitatory HRP fusion constructs) were intersected to increase specificity; (Filter 3) proteins with much greater biotinylation extent by inhibitory constructs than by excitatory constructs were removed on the basis of low 114/115 iTRAQ ratio, to enrich for proteins specific to excitatory synapses (and analogously applied for the inhibitory synapse).

**Figure 3-9.** Histograms that illustrate how Filters 1, 2, and 3 were applied. In filter #1 and #2, green shows the distribution of true positives. Red in filter #1 is the distribution of false positives while grey in filter #2 represents surface proteins that have not been annotated as synaptic. Derivation of iTRAQ ratio cut-off values used is described in the following sections.

*Filtering the data by biotinylation extent (Filter #1)*
To select cutoffs for proteins genuinely biotinylated by HRP and BxxP (Filter #1 Fig. 3-9), we classified the detected proteins into three groups: The first were synaptic cleft exposed proteins (true positive list TP1; these are 176 proteins that have been well characterized and described in the literature as synaptic, often by a combination of electron microscopy, prior synaptosome purifications, fluorescence microscopy, and had a known synaptic function or demonstrated to be synaptogenic which was compiled into a list including references describing their synaptic localization. The second group was intracellular proteins annotated in Gene Ontology (GO) (Ashburner et al., 2000; The Gene Ontology Consortium, 2014) with their cellular component as “nucleus”, “mitochondria”, (and mitochondrial proteins identified in (Hung et al., 2014; Rhee et al., 2013)) “peroxisome”, “lysosome”, “cytosol”, “endoplasmic reticulum” and “golgi” proteins. Since some endoplasmic reticulum and golgi proteins could have surface exposure, proteins annotated as “extracellular”, in TP1, or identified previously in synaptic preparations were removed (Bayés et al., 2012; Biesemann et al., 2014; Boyken et al., 2013; Pirooznia et al., 2012). This list was called false positive list 1 (FP1). We acknowledge that this list may still contain a small number of surface exposed proteins since GO annotation remains incomplete. (3) all other proteins.

ROC, TPR-FPR analysis for log$_2$(LRRTM1/Uninfected)

Figure 3-10. ROC analysis of LRRTM1/uninfected to generate a cutoff for proteins truly biotinylated by LRRTM1. Receiver operating characteristic and true positive rate (TPR) – false positive rate (FPR) analysis to determine optimal iTRAQ cut-off values. Left: For every iTRAQ ratio cut-off, the TPR is plotted against the FPR. Here, the TPR is defined as the fraction of detected true positive synaptic cleft proteins that are above the iTRAQ cut-off. FPR is defined as...
the fraction of detected proteins that are intracellular (FP1) above the SILAC cut-off. Bottom: plots depicting the difference between TPR and FPR at every iTRAQ cut-off.

We selected iTRAQ ratios that maximize the difference between true positive rate (TP1) and false positive rate (FP1) as our cut-offs using receiver operating characteristic (ROC) analysis. A representative plot of one of these analysis is shown in Fig. 3-10. These cut-off values were as follows:

Experiment 1 excitatory cutoffs: 0.875
Experiment 1 inhibitory cutoffs: 0.540
Experiment 2 excitatory cutoffs: 0.774
Experiment 2 inhibitory cutoffs: 0.545
Experiment 3 excitatory cutoffs: 0.862
Experiment 3 inhibitory cutoffs: 0.835

Filtering the data by cleft enrichment (Filter #2)

To select cutoffs for proteins that were selectively biotinylated by HRP targeted to the synaptic cleft versus by general cell surface targeted HRP (Filter #2 Fig. 3-9); truly synaptic proteins, we classified the detected proteins into three groups: (1) synaptic cleft exposed proteins that can be found in TP1 list (available on request), (2) surface exposed proteins identified by searching GO "cellular component"; proteins with one of the following GO annotations: "cell surface", "extracellular space", "extracellular region", "external side of plasma membrane", "extracellular matrix", "extracellular vesicular", "integral component of plasma membrane" AND are not entries containing the keyword "mitochondrion". To reduce the probability of assigning any false negatives to this list, we crossed this list and removed proteins annotated as synaptic (Bayés et al., 2012; Biesemann et al., 2014; Boyken et al., 2013; Pirooznia et al., 2012), and also removed any proteins found in TP1 or FP1. This list is called FP2 and is available on request.

iTRAQ ratio cut-offs for truly synaptic proteins were determined using class (1) and class (2) proteins only. For each experiment and for each iTRAQ ratio value; for the excitatory proteome this was log2 (114/116) while for the
inhibitory proteome this was \( \log_2 (115/116) \), we calculated the true positive rate (TPR) and false positive rate (FPR) we would obtain if we retained only proteins above that iTRAQ ratio for our synaptic clef tome. We defined TPR as the fraction of class (1) proteins (i.e., detected synaptic proteins) above the iTRAQ ratio in question, and FPR as the fraction of class (2) proteins (i.e. general cell surface proteins) above the iTRAQ ratio in question. Although we define class (2) proteins as false positives, in reality this is actually a pseudo false positive list and (2) may indeed have undiscovered genuine synaptic proteins. To compensate for this possibility, we first approximated the probability of finding a non-synaptic protein in a list of surface proteins. This was estimated by taking the total number of proteins in (2), divided by the total number of surface proteins identified in GO "cellular component" (this includes proteins identified in our synaptic protein list or had been identified in a synaptic preparation (Bayés et al., 2012; Biesemann et al., 2014; Boyken et al., 2013; Pirooznia et al., 2012)); 2414/2858 = 0.8446. We applied this as a weight for the FPR when calculating the difference in TPR and FPR; TPR - 0.8446 x FPR, to identify the iTRAQ ratio that maximizes this difference as a cut-off (by ROC analysis). These cutoff values are as follows:

Experiment 1 excitatory cutoffs: 0.435
Experiment 1 inhibitory cutoffs: -0.144
Experiment 2 excitatory cutoffs: 0.443
Experiment 2 inhibitory cutoffs: 0.033
Experiment 3 excitatory cutoffs: 0.366
Experiment 3 inhibitory cutoffs: 0.215

Note that some of these contained multiple local maxima, and we selected the maxima that corresponded to the least stringent cutoff, in order to maximize the coverage in the final proteome.

*Intersection of the proteomic experiments to improve specificity*
After filtering the data based on iTRAQ ratios (Fig. 3-9), we rationalized that an intersection of the 3 enriched protein lists from the three experiments would generate datasets with high specificity. Finally, we performed a last filter based on excitatory versus inhibitory enrichment (114/115 iTRAQ ratio) (Fig. 3-9).

Filtering the data by cleft type enrichment (Filter #3)

To eliminate any false positives from each synaptic clefome following the intersection of 3 experiments, we used a strategy previously employed in our mitochondrial matrix study (Rhee et al., 2013). The log₂ (iTRAQ ratio 114/115) from each of the 3 experiments represents a measure of a protein’s preference for the excitatory or inhibitory synapses, and these ratios were summed and averaged. Separate histograms were plotted for the excitatory or inhibitory synapse proteomes with proteins binned by their log₂(114/115) ratios in increments of 0.1 (Fig. 3-9).

To determine the false positive rate, we first identified a subset of synaptic proteins in TP1 that were known to be exclusive to each synapse type; this was class (1) 54 glutamatergic proteins and class (2) 16 GABAergic proteins. We calculated the false positive rate (FPR) as a function of protein iTRAQ log₂(114/115) ratio using the equation:

\[ FPR \text{ (iTRAQ ratio)} = \frac{P(iTRAQ \text{ ratio} | (2))}{P(iTRAQ \text{ ratio} | (1))} \]

(this applies to the excitatory synaptic clefome, for the inhibitory synaptic clefome, (1) and (2) are swapped)

The numerator is the conditional probability of finding a false positive protein in a particular iTRAQ ratio range and is calculated as the number of false positive proteins (for the excitatory proteome this would be a class (2) protein and class (1) for the inhibitory proteome) detected in that iTRAQ ratio range divided by the total number of false positive proteins detected in the entire
dataset. Likewise, the denominator is the conditional probability of finding a class (1) or class (2) for the excitatory and inhibitory proteomes respectively in a particular iTRAQ ratio range and is calculated similarly. The FPR is the ratio of the two conditional probabilities. Because the synapse type exclusivity of most synaptic proteins was poorly explored, we selected as our cut-off point for each replicate the iTRAQ ratio at which the FPR reaches 0.2. All proteins above this cut-off are >5 times more likely to be specific for the excitatory or inhibitory synapse than the opposing synapse. The cut-offs (in \( \log_2(114/115) \) units) for the excitatory synapse and inhibitory synapse were -0.243 and 0.484, respectively. These cutoffs resulted in an excitatory synaptic cleft proteome of 199 proteins, while giving an inhibitory synaptic cleft proteome of 42 proteins.

<table>
<thead>
<tr>
<th></th>
<th>Experiment 1</th>
<th>Experiment 2</th>
<th>Experiment 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of detected proteins</td>
<td>2790</td>
<td>2436</td>
<td>2696</td>
</tr>
<tr>
<td>Number of biotinylated proteins</td>
<td></td>
<td>615 599 444 428 552 492</td>
<td></td>
</tr>
<tr>
<td>Number of cleft-enriched proteins</td>
<td></td>
<td>422 435 264 171 403 339</td>
<td></td>
</tr>
</tbody>
</table>

**Figure 3-11.** Filtering of MS data to produce excitatory and inhibitory proteomic lists. The table shows the number of proteins remaining after each filtering step. In the first row, a protein was considered detected if two or more unique peptides were sequenced by MS. Filter 1 retains HRP-biotinylated proteins and removes non-specific bead binders, on the basis of the 114/117 or 115/117 iTRAQ ratio. Filter 2 retains cleft-enriched proteins over general cell surface proteins, on the basis of 114/116 or 115/116 iTRAQ ratio. Filter 3 removes strongly inhibitory-enriched proteins (high 115/114 iTRAQ ratio) from the excitatory proteome, and vice versa for the inhibitory proteome.

**Characterization of the excitatory and inhibitory synaptic cleft proteomes**
Figure 3-12. Bar graphs showing the synapse specificity (left) and synapse subtype specificity (middle) of the two proteomic lists. Excit. refers to the excitatory proteome of 199 proteins, and Inhib. refers to the inhibitory cleft proteome of 42 proteins. For synapse subtype analysis, only proteins with literature annotation as excitatory/inhibitory/both are included in the analysis; non-annotated synaptic proteins are excluded. On the far right, proteins are classified according to their sub-synaptic localization.

Determination of surface protein type for proteomic hits

We first verified that no known intracellular proteins were present, consistent with our use of membrane impermeant BxxP for labeling. This was evaluated by identifying proteins containing transmembrane helices using TMHMM (Krogh et al., 2001), a membrane protein topology prediction method. We also searched individual entries in UniProt under “Subcellular location” and “Topology” for evidence of transmembrane annotation, GPI anchor prediction, or if a protein was secreted.

Specificity analysis of synaptic cleftomes

We can evaluate the quality of the 2 proteomes broadly by 3 measures, synapse specificity, synapse subtype specificity and depth of coverage. Synapse specificity refers to how many proteins within the dataset already are annotated in the literature as localized to synapses. The synapse subtype specificity refers to of the synaptic proteins within each proteome, how many are known to already reside, and function within glutamatergic excitatory synapses and GABAergic inhibitory synapses. Lastly the depth of coverage gives us a measure of the
extent by which the proteomes are able to identify bona fide members of each synapse subtype.

**Synapse specificity**

To calculate the synaptic specificity of the excitatory and inhibitory synaptic cleft proteomes (column 1-3 of Fig. 3-12), we first determined if any proteins in the proteomes had prior synapse annotation in GO “cellular component” or were found in TP1. We rationalized that any proteins identified in our proteomes, and previously identified by biochemical fractionations of synapses, would indeed have a high probability of being a synaptic protein, since such a protein would have been identified by two orthogonal methodologies and unlikely to be a biochemical contaminant. Therefore, we crossed our proteomes with rat protein analogs identified by mass spectrometry in the following synaptic preparations: 1) fractionated cortex post-synaptic densities (Bayés et al., 2012) (Table S1 of this reference), 2) purified synaptosomes (Biesemann et al., 2014) (Table S1 and S2 of this reference), 3) presynaptic vesicle docking complexes (Boyken et al., 2013) (Table S1 of this reference), or identified with a reference showing its synaptic localization or function. Additionally we identified any proteins that were identified by Synaptome DB (Pirooznia et al., 2012), a bioinformatics platform for annotating genes comprising the synaptome. Lastly, we searched for the remaining proteins any literature that described their synaptic localization and annotated each one of these individually with its respective reference.

We then divided the number of entries with synapse annotation by the total size of each of the synaptic cleft proteomes (199 proteins for excitatory synapse, and 42 proteins for the inhibitory synapse). For comparison, we also determined the fraction of the entire rat proteome with prior “synapse” annotation (column 1 of Fig. 3-12). To calculate this, we retrieved the entire rat proteome, 38,912 proteins, from the Uniprot-GO Annotation database in July 2015 and identified the number of Uniprot IDs that had the annotation “synapse”; 445, (Only Gene Ontology cellular component terms were retained from this list).
At ~80%, each list is greatly enriched for synaptic proteins compared to the entire rat proteome (1%). The remaining ~20% of proteins (43 total across both excitatory and inhibitory proteomes) with no known connection to synapses could be false positives, or they could be newly discovered synaptic proteins. Our data described in Figs. 3-16 support the latter possibility.

**Synapse subtype specificity**

To characterize synapse subtype specificity, we analyzed the subset of proteins in each list with known excitatory or inhibitory synapse localization (Fig. 3-12, middle). To calculate the synapse type specificity of the excitatory and inhibitory synaptic cleft proteomes (column 4 and 5 of Fig. 3-12), we searched the literature for references identifying each protein in the respective synapse type. Additionally, we crossed our proteomes with proteins identified to be enriched in the vGlut1 sorted synaptosomes (Biesemann et al., 2014), since these are likely to be genuine excitatory synapse proteins. Similarly, we identified proteins that were found in immunoprecipitation experiments of inhibitory synapse proteins as belonging to the inhibitory synapse (Heller et al., 2012; Kang et al., 2014). Note that a caveat of our analysis is that a protein identified as belonging to one particular synapse type may have the possibility of being dual localized, but may not have been studied in the literature and hence be annotated as belonging to only one synapse type. Using these criteria, each protein was placed into one of four categories: excitatory “E”, inhibitory “I”, known to be in either synapse types “both”, or unknown “U”.

The excitatory proteome is highly enriched for proteins known to reside at excitatory synapses or at both synapse types (98%). The inhibitory proteome is enriched for known inhibitory synapse proteins (66%), but also includes a significant number of proteins with excitatory annotation. Because the inhibitory synapse is poorly characterized, and the literature is biased towards assays of excitatory synapse localization (Heller et al., 2012), it is possible that many of these proteins are actually dual-localized to both excitatory and inhibitory
synapses. Indeed, follow-up experiments described below and shown in Figs. 3-17 and 3-18 demonstrate dual localization for at least four excitatory-annotated proteins which appear in our inhibitory cleft proteome.

**Determination of sub-synaptic localization of proteomic hits**

The sub-synaptic localization of each protein in the proteome (column 6 and 7 of Fig. 3-12) was identified using either their GO “cellular component” or UniProtKB “Keywords – Cellular component”, or individual references describing its localization, or inferred based on the sub-synaptic localization given in SynaptomeDB (Pirooznia et al., 2012), or by the sub-synaptic proteome the protein was identified in; if a protein was identified in a post-synaptic density purification (Bayés et al., 2012), we categorize that protein as being found in the post-synapse. Using these criteria, each protein was placed into one of five categories: postsynaptic, presynaptic, both, soluble or unknown. Although our HRP fusion constructs were localized to the post-synaptic membrane, each list contained numerous pre-synaptic proteins, in addition to soluble proteins secreted into the cleft by neurons (e.g., Nptx1 (Xu et al., 2003)) or neighboring glial cells (e.g., Gpc6 (Allen et al., 2012)) (Fig. 3-12). The great majority of detected proteins (83%) were integral membrane proteins, but ~9% were GPI-anchored.

**Depth of coverage analysis of synaptic cleftomes**

*Calculation of depth-of-coverage*

To determine the sensitivity, or depth-of coverage, of each proteome, we generated separate lists of well-established excitatory cleft-resident or inhibitory cleft-resident proteins (available on request). Coverage for each proteome was calculated by the detection of a list of excitatory and inhibitory synapse proteins that had strong support for their localization within each synapse type. These lists are a subset of TP1 (available on request), since not all synaptic cleft proteins
have been distinctively identified as belonging in either excitatory or inhibitory synaptic clefts or both. For example, synaptic proteins without a specific synapse type assignment would include synaptic proteins identified in synaptosomes preparations or identified by electron microscopy.

Of the 62 such excitatory proteins, our proteome contained 43 (69%). Of the 24 inhibitory proteins, we detected 11 (46%). The proteins we failed to detect may be sterically shielded from biotinylation through protein or membrane interactions in the live cell context. Alternatively, they could be dual-localized, with both a synaptic population and extra-synaptic population, and consequently removed by our Filter 2 step (Fig. 3-9), which considers the ratio of biotinylation by synapse-localized HRP versus general cell surface HRP-TM. Finally, some genes may not be expressed in the specific rat cortical neuron preparations that we used.

**Figure 3-13.** Cartoon depicting well-established excitatory (left) and inhibitory (right) synapse proteins. Proteins are colored according to whether they were detected in our excitatory proteome (green), inhibitory proteome (red), both (striped), or neither (grey). Proteins with multiple isoforms are listed below; purple font indicates detection in both proteomes.

Fig. 3-13 provides an illustration of the synapse subtype specificity of both proteomic lists, showing for example that AMPA and NMDA receptor subunits are detected only in the excitatory proteome, while GABA<sub>A</sub> receptor components...
are detected exclusively in the inhibitory proteome. Interestingly, previous studies have suggested that some inhibitory synapse components "leak over" to excitatory synapses, and vice versa, perhaps to enable cross-talk and regulation between synapse types (Chen et al., 2012). Consistent with these studies, our data show that the inhibitory GABA<sub>B</sub> receptor subunit Gabbr1 resides at excitatory synapses as well (Hirono et al., 2001), and the excitatory kainate receptor Grik2 can also be found at inhibitory synapses (Lerma, 2003). 20 proteins were present in both lists, including known dual-localized synaptic proteins such as Erbb4 (Krivosheya et al., 2008), Grik2 (Lerma, 2003), and Gabbr1 (Hirono et al., 2001; Kulik et al., 2002).

![Figure 3-14. Proteins of each proteomic list, subdivided by functional class.](image)

**Figure 3-14. Proteins of each proteomic list, subdivided by functional class.** Genes in blue font have no prior connection to synapses (i.e., they are synapse orphans), while genes with asterisks(*) have no prior connection to that specific synapse type, but are known to be generally synaptic. (D) Scatter plot showing the separation of proteins by E/I (excitatory/inhibitory) ratio. All proteins detected in Experiment 2 are plotted, by biotinylation extent in the inhibitory cleft (y-axis) versus excitatory cleft (x-axis). Each protein is colored according to whether it was detected in either final proteomic list. Points corresponding to some well-established inhibitory (red), excitatory (green), and dual-localized (purple) synaptic proteins are labeled.

**Functional classification of proteins**

Fig. 3-14 shows the functional classification of proteins identified in each proteomic list. In general, the list of synapse orphans can serve as a powerful resource for generation of novel hypotheses. For example, we were intrigued by
the detection of Csmdl at the inhibitory synaptic cleft. Though no literature describes Csmdl as a synaptic protein, the CSMD1 gene has been linked by GWAS studies to schizophrenia (Håvik et al., 2011), and the protein may be part of the complement pathway that facilitates synaptic pruning (Kraus et al., 2006; Stevens et al., 2007). Thus the detection of endogenous Csmdl in the inhibitory cleft of live neurons suggests a possible link between inhibitory synapse elimination and schizophrenia that could be explored in future studies.

Validation of proteomic data

**Novel synaptic proteins identified by imaging**

Within each proteome, ~80% of proteins have prior literature connection to synapses. The remaining ~20% are “synapse orphans”, with no previous literature assigning them to synapses. We found 36 such orphans in the excitatory cleft, 3 in the inhibitory cleft, and 4 in both clefts (some examples in blue font in Fig. 3-14; complete list of orphans is available on request). We performed follow-up fluorescence imaging on several of these synapse orphans, to determine if they were either false positives or novel synaptic proteins. Many genes were intractable to study because they were too large to clone into lentiviral vectors (>5000bp such as CSMD1/3) and most of the orphans did not have specific antibodies available against the endogenous proteins suitable for fluorescence microscopy validation.
Figure 3-15. The HRP tag is superior to Venus fluorescent protein for detection of surface proteins at low expression levels. (A) BA13 constructs compared here. (B) Venus and HRP fusions to BA13 were introduced to cultured neurons in three expression level regimes: highest, via lipofectamine transfection; lower, via lentiviral transduction; and lowest, via lentiviral transduction but with 100-fold less virus. The last column is the same as the third column, but the intensity scale is narrowed to show the images at higher contrast. At DIV19, neurons were labeled live with BxxP, then fixed without permeabilization and stained with neutravidin-AlexaFluor647 to detect BxxP labeling and anti-V5 antibody (AlexaFluor568 readout) to detect the V5 tag. Whereas HRP-BA13 puncta are visible, Venus-BA13 puncta cannot be detected at the lowest expression levels. (C) Same as B, except the anti-V5 staining was performed after cell permeabilization, in order to detect total protein pools rather than cell surface pools only.

Because of these challenges, we explored a number of tagging strategies, either by epitope tagging, fluorescent protein tagging or with our HRP construct. We found that HRP dependent BxxP labeling was more sensitive than GFP and epitope staining for detection of orphan localization (because of signal amplification), and many of these proteins had to be expressed at very low levels to visualize synaptic localization since we are exogenously introducing these proteins on top of their endogenous pool.

Nonetheless we tagged 10 orphan genes with HRP, GFP and V5 tags and found that six out of six integral membrane proteins we tested from the excitatory proteomic list showed clear localization to excitatory synapses (Cxadr, Pcdh9, Bai3, and Bmpr2, Egfr and Lingo2; Figs. 3-16). Low expression levels was
required to observe punctate localization for these proteins, and in most cases this required utilizing an HRP tag (with BxxP labeling following by streptavidin staining) for highly sensitive and amplified visualization, since GFP and V5 tags were not detectable at these expression levels. We also performed fluorescence imaging on the three secreted synapse orphans in our excitatory proteomic list (Smpdl3b, Brinp2, and Brinp3). For all, we observed numerous puncta but no clear overlap with synaptic markers (data not shown). These proteins could be false positives, or perhaps recombinant fusion to HRP disrupts their secretion and localization to the cleft space. In support of the latter possibility, Brinp1, which is related to Brinp2 and 3, has previously been identified in synaptosome preparations (Biesemann et al., 2014).

For the remaining synapse orphans that we did not test by microscopy, several have indirect evidence linking them to synapses. For example, Pcdh19 binds to known synapse protein N-cadherin (Biswas et al., 2010).

Figure 3-16. Fluorescence imaging of four synapse orphans. (A) Colocalization of five synapse orphans with pre-synaptic marker Bassoon. Orphan genes were fused at their N-terminal ends to HRP, and expressed via lentiviral transduction in DIV19 dissociated rat cortical neurons. The HRP tag was visualized by live BxxP labeling, followed by neutravidin-AlexaFluor647 staining on fixed cells. Endogenous Bassoon was detected with an antibody followed by AlexaFluor568 readout. HRP-LRRTM2 is a positive control, and GFP is a non-
synapse localized negative control. (B) Quantitation of data in (A). >8 fields of view were analyzed for each sample. Errors, ±1 s.e.m. (C) Colocalization of orphans in (A) with excitatory synapse marker vGlut1. Samples were prepared and stained as in (A), and endogenous vGlut was detected by antibody staining with AlexaFluor568 readout.

**Synapse subtype assignment of synaptic proteins and validation**

Apart from revealing novel synaptic proteins, our datasets provide new or further localization information for proteins already known to be synaptic. For example, many proteins have been identified in synaptosome preparations, but it is unclear whether they are excitatory-specific, inhibitory-specific, or both. Other proteins have been classified in the literature as excitatory synaptic proteins, but our data newly reveals an inhibitory synapse-localized population as well. In aggregate, our proteomes assign 27 of these “synapse subtype orphans” to the excitatory synaptic cleft, 12 to the inhibitory synaptic cleft, and 2 to both synapse types (some examples are starred in Fig 3-14; complete list available on request).

For follow-up validation by fluorescence microscopy, we selected five of the most surprising synapse subtype orphans: those known in the literature as excitatory synapse-specific proteins that we unexpectedly also detected in our inhibitory cleft proteome (Fig. 3-17 and 3-18). We generated HRP or GFP fusions to each of these proteins (Flt2, Elfn1, Ephb6, Dcc, and Nlgn1), and imaged them at low expression levels together with excitatory (anti-vGlut1) and inhibitory (anti-vGAT) synapse markers. Consistent with our proteomic data, Figs. 3-17 and 3-18, Fig. 2-14 show that all five proteins are localized to both excitatory and inhibitory synapses. The biological implications of these findings are quite interesting and vary for each protein.
ELFN1-GFP (view showing excitatory preference) ELFN1-GFP (view showing inhibitory preference)

Figure 3-17. Imaging of synapse subtype orphan Elfn1, which appeared in both our proteomic lists, despite being classified as an excitatory-specific protein in the literature (Sylwestrak and Ghosh, 2012). ELFN1-GFP was expressed in DIV19 cultures via lentiviral transduction, and imaging was performed after fixation and staining with anti-vGAT and anti-vGlut1 antibodies. Within the same culture dish, Elfn1 appeared at only excitatory synapses in some neurons (left view), and at only inhibitory synapses in other neurons (right view).

Figure 3-18. Imaging of synapse subtype orphans Flrt2, Ephb6, and Dcc along with synapse orphan Pcdh9. Neurons expressing HRP fusions to these genes were prepared and stained as in (B). All four proteins show overlap with both excitatory and inhibitory synapse markers.

For example, Elfn1, a trans-synaptic binding partner of presynaptic mGlur7 (Tomioka et al., 2014), was previously shown to localize specifically to excitatory post-synapses in O-LM interneurons (Sylwestrak and Ghosh, 2012). Because we detected endogenous Elfn1 in both our proteomes, we proceeded to image recombinant Elfn1-GFP at low expression level in cultured cortical neurons. Interestingly, in some neurons in these cultures, we observed Elfn1
localization to excitatory synapses only, while in other neurons within the same culture, Elfn1 was inhibitory synapse-specific (Fig. 3-17). This observation suggests that alternative mechanisms may exist within different neuron types to regulate the post-synaptic localization of Elfn1. At inhibitory synapses, we speculate that Elfn1 may be involved in a different form of mGluR-related signaling, perhaps via trans binding to mGluR4 (71% homology to mGluR7) that we also detect in our inhibitory cleft proteome. The biological context and possible implications of the four other proteins we discovered to be dual-localized to both excitatory and inhibitory synapses (Flr2, Ephb6, Nlgn1, and Dcc) are discussed below.

Among the remaining proteins in the list of 40 synapse subtype orphans that we did not validate by fluorescence microscopy, other interesting observations may be made as well. Neurexins are pre-synaptic transmembrane adhesion proteins that bind to post-synaptic neuroligins (Ichtchenko et al., 1995, 1996) and Lrrtms (Ko et al., 2009; Siddiqui et al., 2010). Because available antibodies do not distinguish between neurexins-1-3, it has not been possible to determine if their endogenous localizations are the same or distinct. Our proteomic data shows that neurexin-1 is preferentially localized to the excitatory cleft, neurexin-3 is in the inhibitory cleft, and neurexin-2 can be found at both cleft types. Although the lack of detection of neurexin isoforms in each synaptic cleft type does not preclude it from being in those cleft types (due to our low coverage), it does suggest that they may be enriched and have functions related to those synapse types.

**Biological Implications of new assignments**

Several of the new orphans are known to bind to synaptic proteins. For example, Egfr binds to known synapse protein Erbb4 (Tao and Maruyama, 2008), while Cxadr is known to possess a PDZ binding domain and binds to PSD95 in non-neuronal cells (Excoffon et al., 2004).
Reelin is another synapse subtype orphan that caught our attention. Though we detected it in both proteomes, it is known in the literature as a secreted glycoprotein localized to excitatory synapses, based on its interaction with apolipoprotein E receptor Apoer2 at excitatory terminals (Beffert et al., 2005; Hiesberger et al., 1999), and its ability to regulate NMDA receptor function there (Chen et al., 2005; Sinagra et al., 2005). Reelin also binds to Vldlr (Hiesberger et al., 1999) and regulates neuron migration and lamination of the cortex (Trommsdorff et al., 1999). Interestingly, Reelin is known to be secreted by GABAergic interneurons (Pappas et al.), and if our proteomic data are accurate, Reelin may spend some time in the inhibitory synaptic cleft before or in addition to localizing to the excitatory cleft. We speculate that Reelin may have a function in crosstalk between synapse types.

Flrt2 is a member of the family of leucine rich proteins (Flrt) that are post-synaptic ligands of the latrophilins (Lphn). Flrt2 and Flrt3 are known to bind Lphn3 \textit{in vitro} (O’Sullivan et al., 2012), and the Lphn3-Flrt3 complex regulates excitatory synapse numbers in cultured neurons (O’Sullivan et al., 2012). However, Flrt2’s function at synapses is unexplored. Flrt3 and Lphn3 were both detected in only our excitatory proteome while Flrt2 was detected in both proteomes. Does Flrt2 have a similar function to Flrt3 or different? Is the function of Flrt2 at excitatory synapses the same or different from its function at inhibitory synapses, where latrophilins are apparently absent? Interestingly, the Flrts are competitive with teneurins as ligands for latrophilin (Boucard et al., 2014), and teneurins were also detected in both our cleft proteomes. None of these proteins have previous links to the inhibitory synapse, and their presence there (revealed by our proteomes) opens up new questions about the scope of their biological roles.

Ephb6 is another known excitatory synapse protein that we found (by both proteomics and imaging) to be also localized to inhibitory synapses. Ephrin signaling plays a role in excitatory synapse formation and function (Hruska and Dalva, 2012), but has not previously been linked to inhibitory synapses. Ephb6 is a catalytically inactive ephrin receptor that may heterodimerize with other ephrin
receptors (Freywald et al., 2002). Is Ephb6 functioning in ephrin-related signaling at inhibitory synapses, or does it have a completely different function there?

Nlg1 is well-characterized as an excitatory synapse-specific post-synaptic adhesion protein, yet we detected in both our excitatory and inhibitory proteomes. Fluorescence imaging in Fig 2-41 showed that it overlaps with both excitatory and inhibitory markers, in contrast to Nlg2, which is inhibitory-specific. Functionally, Nlg1 is a potentiator of excitatory synapses (Chubykin et al., 2007). Perhaps it localizes to inhibitory synapses via heterodimerization with Nlg2 (Poulopoulos et al., 2012), or via binding to the inhibitory scaffold protein gephrin, which recognizes Nlg1’s cytosolic domain based on its pattern of phosphorylation (Giannone et al., 2013). Further support for an inhibitory synapse pool of Nlg1 is provided by the preliminary proteomic data that we obtained using an HRP-NLGN1 fusion construct (Fig. 2-35). Numerous endogenous inhibitory synapse proteins, such as GABA receptors and Igsf9b were biotinylated and enriched by this construct, in addition to the expected excitatory synapse markers.

Finally, Dcc is a netrin receptor involved in axon guidance (Stein and Tessier-Lavigne, 2001) and synapse formation (Goldman et al., 2013) that we unexpectedly detected in our inhibitory synaptic cleft proteome, and confirmed by fluorescence microscopy in Figure 3-18. Although previously shown to be located at excitatory synapses only in cultured mouse neurons (Goldman et al., 2013), Dcc has recently been linked to clustering of GABA receptors via the MADD4 protein in C. elegans neuromuscular junctions (Maro et al., 2015; Tu et al., 2015). An interesting question for future investigation is whether the mammalian homolog of MADD4 (ADAMTSL1/3 (Pinan-Lucarré et al., 2014)) is also involved in GABA receptor regulation at inhibitory synapses, via the action of Dcc.

Protein isoforms with distinct localizations
Due to our experimental design, every protein appearing in our lists is associated with an E/I (excitatory/inhibitory) ratio, based on the 114/115 iTRAQ ratio, that reflects its enrichment at excitatory versus inhibitory synapses. This can be visualized in the scatter plot shown in Fig. 3-19. Known excitatory synapse-specific proteins such as AMPA receptors (Gria2-3) and NMDA receptors (Grin1) appear below the diagonal, whereas known inhibitory synapse-specific proteins such as GABA receptors (Gabral,3, Gabrbl-3) lie above the diagonal. Known dual localized proteins such as Erbb4 are close to the diagonal. Below, we have used this quantitation to formulate hypotheses regarding the synapse subtype specificities of various proteins.

While many post-synaptic membrane proteins exhibit synapse subtype specificity (e.g., AMPA, NMDA, and GABA receptors), there are far fewer examples of presynaptic proteins with preferential localization to excitatory or inhibitory synaptic clefts (Boyken et al., 2013). To our knowledge, Caspr4 is the only example of a validated presynaptic membrane protein with specificity for inhibitory over excitatory synapses (Karayannis et al., 2014). As described above, our proteome identified neurexin-3 as another possible such inhibitory-specific presynaptic transmembrane protein. We also observed presynaptic...
Cacna2d2, a protein associated with voltage gated calcium channels (Barclay et al., 2001; Witcher et al., 1993), exclusively in our inhibitory proteome, while the related Cacna2d1 subunit was detected exclusively in our excitatory proteome. Cacna2d3 was detected in both proteomes. Perhaps these proteins have unexplored roles in establishing or maintaining the specificity of pre-synaptic inhibitory and excitatory terminals. Interestingly, separate studies suggest that the expression of Cacna2d2 and Cacna2d1 may be correlated with inhibitory and excitatory neuron types, respectively (Barclay et al., 2001; Cole et al., 2005), consistent with our proteomic observations. Following this train of thought, we decided to follow up on functional studies of a family of proteins, the MDGAs that will be discussed in the Chapter 4.

Conclusions

Microscopy and MS-based proteomics have complementary strengths: preservation of spatial information on the one hand, and unbiased, high-throughout detection of endogenous proteins on the other hand. Techniques like APEX and BiolD have the potential to merge the strengths of these complementary approaches. However, existing demonstrations have not fully explored the potential of proximity biotinylation to chart the molecular composition of important, cellular nanodomains with minute-long temporal resolution and spatial resolution approaching the dimensions of single proteins.

We have used peroxidase mediated biotinylation to map the proteomes of the excitatory and inhibitory synaptic clefts, two subcellular regions which previously could not be isolated for proteomic identification of its components. These proteomic maps were highly specific, reveal dozens of new synaptic proteins, and helped elucidate the similarities and differences between excitatory and inhibitory synaptic junctions that mediate action potential propagation versus inhibition.
Much of our current understanding of the molecular architecture of synapses is based on labor-intensive one-by-one identification and study of individual proteins. Increasingly, this knowledge has been supplemented by large-scale purification-based proteomic studies. Synaptosome preparations, which include both pre- and post-synaptic material along with the cleft, provide a rich source of novel synaptic protein candidates, but do not distinguish between excitatory and inhibitory and other synapse types, and require extensive follow-up experimentation to identify *bona fide* cleft-resident proteins due to the high rate of false positives (Biesemann et al., 2014). Post-synaptic density purifications are more focused (Bayés et al., 2012), but also contain much more than just cleft material, and are not applicable to inhibitory synapses, which lack post-synaptic densities. The Craig (Kang et al., 2014) and Heintz (Heller et al., 2012) labs have used the more focused approach of immunoprecipitation (of NLGN2 and GABA\(_{\alpha}R\)\(\alpha1\), respectively) to identify novel proteins at inhibitory synapses. While these interactomes can be quite specific (Heller et al., 2012), the majority of proteins resident in the inhibitory cleft are missed because they do not interact directly and/or stably with the bait proteins. We overcame these challenges by using a fundamentally different approach, in which very specific protein neighborhoods within living, intact neurons are labeled, resulting in the capture and identification of not only direct interaction partners but also proximal proteins that do not directly touch our “bait” (the HRP fusion construct).

This study was performed on cultured neurons, which are heterogeneous, have synapses of varying maturity, and contain very few astrocytes and other glial cells. The full power of peroxidase-based proteomics will only be realized *in vivo*, where specific neuron types, specific circuits, neuron-glial interactions and other types of synapses (e.g., cholinergic, dopaminergic) could potentially be examined and we hope to develop and take this technology into a completely *in vivo* setting in the future.

Nonetheless, the two proteomes can be mined for biological insights and serve as rich resources for the generation of novel hypotheses. Potentially 34 novel synaptic proteins have been highlighted by our datasets, of which we
validated six by microscopy (Fig. 3-16). Because of the experimental design, every detected protein is associated with an E/I (excitatory/inhibitory) ratio, which reflects the extent to which it is enriched at excitatory or inhibitory synapses. Such quantitation enabled us to clarify the synapse sub-type specificity of dozens of synaptic proteins, including several that had been known in the literature as excitatory-specific proteins, but we discovered an inhibitory synapse-localized population as well (e.g., Flrt2, Ephb6, Elfn1, Dcc, and Nlgn1, all of which we confirmed by microscopy). These observations open up fascinating questions regarding the mechanisms of their localization to inhibitory synapses and their possible biological functions there.

Materials and Methods

Proteomic mapping of the synaptic cleft

Biotinylation in neurons and streptavidin enrichment of biotinylated proteome

Cortical neurons were cultured as described under “Rat cortical neuron culture”. 2.5 million cortical neurons were plated on each laminin- and poly-D-lysine-coated 10 cm dish, and six dishes were used per iTRAQ channel. There was a total of 24 dishes of cortical neurons used per experiment, or 72 dishes for 3 independent experiments. Cortical neurons at DIV 15 were infected with lentiviruses expressing the HRP-tagged synaptic constructs at a previously optimized titer (see “Titrating lentiviral constructs”) that gives tight, localized synaptic BxxP labeling. To uniformly infect the cells, viruses for each channel were first diluted in 6 ml of Neurobasal, and then evenly distributed to each of the six dishes.

Cells were labeled at DIV 19 with 100 μM BxxP and 1 mM H₂O₂ in Tyrode’s Buffer for 1 min at room temperature, and then quenched with 4 washes of
Tyrode’s buffer containing 10 mM sodium azide, 10 mM sodium ascorbate, and 5 mM Trolox. After quenching, the cells were harvested by scraping and pelleted by centrifugation at 3,000 g for 10 minutes. The supernatant was discarded and the pellet was stored at −80 °C overnight.

The cell pellet was lysed by incubating in 400 ul of 1% SDS lysis buffer (50 mM Tris-HCl, 1% SDS, pH = 8.0) containing protease inhibitor cocktail (Sigma Aldrich, catalog no. P8849), 1 mM PMSF (phenylmethylsulfonyl fluoride), 10 mM sodium azide, 10 mM sodium ascorbate, and 5 mM Trolox, then boiled for 5 min to denature and separate postsynaptic density proteins from each other. This was then diluted in 1600ul of 1.25X RIPA lysis buffer (50 mM Tris-HCl, 187.5 mM NaCl, 0.625% sodium deoxycholate, 1.25% Triton X-100) to give 1X RIPA lysis buffer (50 mM Tris-HCl, pH 8.0, 150 mM NaCl, 0.2% SDS, 0.5% sodium deoxycholate, 1% Triton X-100), Lysates were centrifuged at 16,000 g for 10 minutes at 4 °C. Protein concentrations of the supernatant were measured using the Pierce 660 nm Protein Assay kit, with bovine serum albumin as the reference standard.

Streptavidin-coated magnetic beads (Pierce catalog no. 88817) were prepared by washing twice with RIPA lysis buffer. For each experiment, an equal amount of protein per condition (7.0 mg of protein), was incubated with 450 µl of streptavidin bead slurry overnight at 4 °C with gentle rotation. The beads were washed with 2 × 1 ml RIPA lysis buffer, 1 x 1 ml of 1 M KCl, 1x 1 ml of 0.1 M Na₂CO₃, 1 x 1 ml of 2 M urea in 10 mM Tris-HCl, pH 8.0, and again with 2 × 1 ml RIPA lysis buffer. Finally the beads were resuspended in 1 ml of RIPA lysis buffer before on bead digestion.

On-bead trypsin digestion

Proteins bound to streptavidin beads were washed 2x with 200 ul of 50 mM Tris HCl buffer (pH 7.5) followed by 2x washes with 2 M urea/50 mM Tris (pH 7.5)
buffer. The final volume of 2 M urea/50 mM Tris (pH 7.5) buffer was removed and beads were incubated with 80 ul of 2 M urea/50 mM Tris buffer containing 1 mM DTT and 0.4 ug trypsin. Beads are incubated in the urea/trypsin buffer for 1 hr at 25°C while shaking (1000 rpm). After 1 hr, the supernatant was removed transferred to a fresh tube. The streptavidin beads were washed 2x with 60 ul of 2 M urea/50 mM Tris (pH 7.5) buffer and the washes were combined with the on-bead digest supernatant. The eluate was reduced with 4 mM DTT for 30 min at 25°C with shaking (1000 rpm). The samples were alkylated with 10 mM iodoacetamide and incubated for 45 min in the dark at 25°C while shaking (1000 rpm). An additional 0.5 μg of trypsin was added to the sample and the digestion was completed overnight at 25°C with shaking (700 rpm). After overnight digestion, the sample was acidified (pH < 3) by adding formic acid (FA) such that the sample contained ~1% FA. Samples were desalted on C18 stage tips and evaporated to dryness in a vacuum concentrator as previously described (Hung et al., 2014).

iTRAQ labeling of peptides

Desalted peptides were labeled with iTRAQ (4-plex) reagents as directed by the manufacturer (Sciex, Foster City, CA). Peptides were resuspended in 30 ul dissolution buffer and 70 ul ethanol. One unit of iTRAQ labeling reagent was used for each condition in a given iTRAQ 4-plex cassette and samples were labeled using the scheme described in Fig. 3-3. Samples were incubated with iTRAQ reagents for 1 hr at RT. iTRAQ labeling reactions were quenched with 10 ul of 1 M Tris HCl pH 8. Differentially labeled peptides were mixed to generate 4-plex iTRAQ samples, desalted on C18 Stage tips and evaporated to dryness in a vacuum concentrator.

SCX stage tip fractionation of peptides
For each iTRAQ 4-plex cassette, 50% of the sample was fractionated by Strong Cation Exchange (SCX) using StageTips (Rappsilber et al., 2007) while the other 50% of each sample was reserved for LC-MS analysis by a single-shot, long gradient. One SCX StageTip was prepared per sample using 3 plugs of SCX material (3M, #2251) topped with 2 plugs of C18 material (3M, #2215). StageTips were washed with 100 µl methanol, then with 100 µl 80% acetonitrile/0.5% acetic acid, and equilibrated with 100 µl 0.5% acetic acid. Samples were reconstituted in 0.5% acetic acid, loaded onto the StageTip and then transeluted from the C18 discs to the SCX discs using 100 µl 80% acetonitrile/0.5% acetic acid. Three step-wise elutions from the SCX disks were completed as follows: the first fraction was eluted with 50 µl of 50 mM NH₄AcO; 20% MeCN (pH 5.15, adjusted with acetic acid), the second with 50 µl 50 mM NH₄AcO; 20% MeCN (pH 8.25, adjusted with acetic acid), the third with 50 µl 50 mM NH₄AcO; 20% MeCN (pH 10.3, adjusted with acetic acid). Each eluate was collected separately and 200 µl of 0.5% acetic acid was added to each. Each fraction was desalted on C18 StageTips and evaporated to dryness in a vacuum concentrator.

**Liquid chromatography and mass spectrometry**

Desalted peptides were resuspended in 9 µl of 3% MeCN, 0.1% FA and analyzed by online nanoflow liquid chromatography tandem mass spectrometry (LC-MS/MS) using a Q Exactive mass spectrometer (Thermo Fisher Scientific) coupled on-line to a Proxeon Easy-nLC 1000 (Thermo Fisher Scientific). Four microliters of each sample was loaded at 500 nl/min onto a microcapillary column (360 µm outer diameter × 75 µm inner diameter) containing an integrated electrospray emitter tip (10 µm), packed to approximately 24 cm with ReproSil-Pur C18-AQ 1.9 µm beads (Dr. Maisch GmbH) and heated to 50 °C. The HPLC solvent A was 3% MeCN, 0.1% FA, and the solvent B was 90% MeCN, 0.1% FA. Peptides were eluted into the mass spectrometer at a flow rate of 200 nl/min. SCX fractionated samples were run using a 150 min inject-to-inject LC-MS method, using the gradient previously described (Svinkina et al., 2015).
single-shot, long runs a 290 min inject-to-inject LC-MS/MS method was utilized. For the 290 min method, after a 1 min ramp to 6% B, a gradient of 0.1% B/min was applied for 234 min followed by a ramp to 60% B (3.3% B/min). The Q Exactive was operated in the data-dependent mode acquiring HCD MS/MS scans (r = 17,500) after each MS1 scan (r = 70,000) on the top 12 most abundant ions using an MS1 target of $3 \times 10^6$ and an MS2 target of $5 \times 10^4$. The maximum ion time utilized for MS/MS scans was 120 ms; the HCD-normalized collision energy was set to 28; the dynamic exclusion time was set to 20 s, and the peptide match and isotope exclusion functions were enabled. Charge exclusion was enabled for charge states that were unassigned, 1 and >7.

**Analysis of proteomic data**

Working up the raw mass spectrometric data (includes SpectrumMill searching, unique peptide filter, and normalization)

All MS data were interpreted using the Spectrum Mill software package v5.0 prerelease (Agilent Technologies, Santa Clara, CA). Similar MS/MS spectra acquired on the same precursor m/z within +/- 60 sec were merged. MS/MS spectra were excluded from searching if they failed the quality filter by not having a sequence tag length > 0 or did not have a precursor MH$^+$ in the range of 750-4000. MS/MS spectra were searched against a UniProt database containing 29,055 rat proteins and <200 common laboratory contaminants. All spectra were allowed +/- 20 ppm mass tolerance for precursor and product ions, 30% minimum matched peak intensity, and trypsin allow P enzyme specificity with up to 4 missed cleavages. Carbamidomethylation at cysteine and iTRAQ at N-termini and lysine were fixed modifications. Allowed variable modifications were oxidized methionine and N-terminal protein acetylation. Individual spectra were automatically designated as confidently assigned using the Spectrum Mill autovalidation module. Specifically, a target-decoy based false-discovery rate (FDR) scoring threshold criteria via a two-step auto threshold strategy at the
spectral and protein levels was used. First, peptide mode was set to allow automatic variable range precursor mass filtering with score thresholds optimized to yield a spectral level FDR of <1.2%. A protein polishing autovalidation was applied to further filter the peptide spectrum matches using a target protein-level FDR threshold of 0.

Each of the three iTRAQ experiments was analyzed separately. Contaminants and proteins identified as reverse hits were removed. For each experiment, only proteins with two or more unique quantified peptides were considered “detected” and retained for further analysis. Unique peptides are those that are designated by SpectrumMill software as peptides not shared with other protein groups.

To normalize the proteomic data, we first identified proteins that could not be biotinylated by extracellular HRP and BxxP; i.e. rat homologs of mitochondrial proteins. This list was compiled using mitochondrial proteins identified from the mitochondrial matrix proteome (Rhee et al., 2013) the inner membrane space (IMS) of the mitochondria (Hung et al., 2014) and proteins identified in Gene Ontology “Cellular Component” as associated with the mitochondria. Protein iTRAQ ratios were normalized (after protein filtering based on the unique peptide count as described above) so that the median log$_2$ ratio of the distribution of mitochondrial proteins identified was 1.

**Imaging of HRP fusions to synapse orphans**

Cortical neurons were cultured as described under “Rat cortical neuron culture”. Cortical neurons at days in vitro (DIV) 15 were infected with lentiviruses expressing the either the GFP/HRP-tagged synaptic constructs at an MOI less than 1 to infect only a subset of the neurons present in the coverslip.

For HRP tagged synaptic constructs, cells were labeled at DIV 19 with 100 μM BxxP and 1 mM H$_2$O$_2$ in Tyrode’s Buffer for 1 min at room temperature, and then
quenched with 3 washes of Tyrode’s buffer containing 10 mM sodium azide, 10 mM sodium ascorbate, and 5 mM Trolox. The samples were then fixed with 4% paraformaldehyde in “fixation buffer” at room temperature for 10 minutes. The cells were then washed three times with DPBS.

Synaptic localization was inferred from the BxxP labeling. The sample was blocked at room temperature with 3% w/v BSA in DPBS for one hour at room temperature or overnight at 4 °C, then stained with neutravidin protein (Invitrogen) pre-coupled to AF647-NHS esters (Invitrogen), at a 1:1000 dilution for one hour. Samples were then washed three times with DPBS, then permeabilized as described above for staining with antibodies to endogenous vGlut1 and vGAT. Following staining, samples were then imaged by confocal microscopy using a 100x objective.

For GFP tagged synaptic constructs, cells were fixed at DIV 19 with 4% paraformaldehyde in “fixation buffer” at room temperature for 10 minutes. The cells were then washed three times with DPBS. Samples were permeabilized as described above for staining with antibodies to endogenous vGlut1 and vGAT. Following staining, samples were then imaged by confocal microscopy using a 100x objective.

References


Chapter 4: Identification of MDGA2 as an inhibitory synapse specificity factor

Sections of this chapter are excerpts taken from a submitted manuscript: K. H. Loh, P. S. Stawski, N. D. Udeshi, T. Svinkina, T. Deerinck, M. H. Ellisman, S. A. Carr, and A. Y. Ting. Proteomic maps of the excitatory and inhibitory synaptic clefts reveal specificity factor for pre-to-post synaptic matching. Electron microscopy and knockout experiments were kindly performed in collaboration with Philipp Stawski.
Introduction

An open and unsolved problem in molecular neuroscience pertains to the molecular mechanisms by which accurate matching is achieved between neurotransmitters in the presynapse and their cognate receptors in the postsynapse, during the process of synapse formation. Inspired by the neuroligin family of proteins, which have different isoforms that localize to functionally divergent synapse types (ie. Nlgn1 localizes to excitatory glutamatergic synapses (Song et al., 1999), while Nlgn2 localizes to GABAergic inhibitory synapses (Poulopoulos et al., 2009; Varoqueaux et al., 2004), and each of them has a functional role belonging to that synapse type. We decided to explore the two functionally different synaptic cleft proteomes, to identify protein families where the function of the proteins within that family were approximated to be similar (by virtue of belonging to the same family), but had differential localizations to the different synapse types, because 1) this would allow us to generate a comparative analysis of the protein isoforms and identify the domains responsible for this localization, and 2) possible clues to their functional roles in that synapse type, since we could use the already studied isoforms as a template for inferring what each isoform might be doing. We chose to look at adhesion proteins because these were most likely to participate in a “synaptic code” to try to get closer to answering the molecular mechanisms by which accurate matching is achieved. Inspection of our proteomes pointed us to a number of candidates that fit these criteria, such as the Cacna2d family of calcium channel associated proteins, the neurexins and the one we decided to follow up on was a family of proteins called the MDGAs.

MDGA domain structure

The MDGAs (MAM domain-containing glycosylphosphatidylinositol anchor) are a family of GPI anchored proteins that have previously been implicated in autism spectrum disorders and schizophrenia (Kähler et al., 2008;
Li et al., 2011). There are two members, MDGA1 and MDGA2, that are highly expressed within the brain and these proteins possess the following domain structure; each MDGA has six Ig domains, a fibronectin type III repeat (FNIII), a single meprin/A5 protein/receptor protein tyrosine phosphatase mu (MAM) domain, followed by a C-terminal GPI anchor. These two members are highly homologous (~70% by identity) and are expected to mediate adhesion between neurons based on the presence of their Ig domains, much like other members of the Ig domain superfamily members; such as the neural cell adhesion molecule (NCAM) and L1 cell adhesion molecule (L1-CAM), two known synaptic proteins also identified in our synaptic cleft proteomes.

**MDGAs regulates inhibitory synapse formation**

During early neural development, the MDGAs contribute to the radial migration and positioning of a subset of cortical neurons (Takeuchi and O'Leary, 2006). The MDGAs continue to be expressed postnatally, and were discovered to play a role in inhibitory synapses via their binding of the inhibitory post-synapse organizing adhesion protein, neuroligin 2. *In vitro* cell surface binding assays (Lee et al., 2013; Pettem et al., 2013) showed that MDGA1 and 2 were found to bind selectively and directly to the purified ectodomain of neuroligin 2 fused to an Fc domain with nanomolar $K_d$ (7.3 ± 1.0 nM for Nlg2-Fc binding to MDGA1 and 45.9 ± 11.9 nM for MDGA2). In both studies, after demonstrating that MDGA1 and MDGA2 both bound to neuroligin 2 in vitro, the authors focused their assays on MDGA1 because of its higher expression levels in the brain and slightly higher binding affinities to neuroligin 2, and made the assumption that they would be similar in function and localization subcellularly. MDGA1 was found to have the interesting function of negatively regulating inhibitory synapse formation; it is able to inhibit the synaptogenic activity of neuroligin 2 by suppressing the binding of neuroligin 2 to its trans-synaptic partner neurexin, an interaction that is important for synapse formation. MDGA1 knockdown also causes an increase in inhibitory synapse density in cultured neurons, and this
effect is abolished when neuroligin 2 is knocked out. Conversely, overexpression of MDGA1 results in a decrease of inhibitory synapses. This function is analogous to the Nogo receptor complex that regulates excitatory synapse density (Wills et al., 2012), a complex that we also identified in the excitatory cleft proteome.

Despite the assumed similarities between MDGA1 and MDGA2 in terms of its function and localizations, we were surprised to find MDGA1 in our excitatory cleft proteome, and MDGA2 within our inhibitory cleft proteome (Fig. 4-1).

**Figure. 4-1. Proteomic Identification of the MDGA family in the synaptic cleft proteomes.** MDGA1 and MDGA2 were differential localized in the excitatory and inhibitory synaptic cleft proteomes respectively.

**MDGAs are localized at synapses**

Because there are no specific antibodies to each MDGA, we first validated their localizations by tagging the MDGAs at the N terminus with HRP, and introduced these into cultured neurons by lentiviral induction at low MOIs, visualizing their localizations either by electron microscopy, or by BxxP labeling.

We first demonstrated localization of MDGA1 and 2 at synapses, at the submicron level by electron microscopy (see Fig. 4-2). MDGA localization was next validated by BxxP labeling, which we could use to identify their synapse sub-type localization.
We observed an exclusive localization of recombinant MDGA2 to inhibitory synapses as visualized by the colocalization with vGAT, and almost no colocalization with the excitatory synapse marker vGlut1 (Fig. 4-3). Recombinant MDGA1 is preferentially localized at excitatory synapses, and at inhibitory synapses too, a result consistent with results previously seen by the Craig lab (Pettem et al., 2013). In that study, the Craig lab demonstrated that when GFP tagged recombinant Mdga1 is expressed in culture neurons, it localizes to both inhibitory and excitatory post-synaptic membranes (Pettem et al., 2013), based on its colocalization with the post-synaptic inhibitory and excitatory markers gephyrin and PSD95. Since our depth of coverage for the inhibitory cleft proteome is not high, it is plausible that we fail to identify it in our inhibitory synaptic cleft proteomic experiment, or that this additional localization to inhibitory synapses might come from recombinant or overexpression artifacts. Nonetheless because the fluorescence microscopy localization patterns of the two proteins were different, echoing the proteomic data, we rationalized that there was definitely a difference between the MDGA isoforms.

Figure. 4-2. EM images showing synaptic localization of HRP fusions to Mdga1 (left) and Mdga2 (right). Samples were prepared and stained as in Fig. 2-44 in Chapter 2. Scale bars, 200 nm.
Synaptogenesis assay to identify and compare MDGA function

The differential localization prompted us to wonder if the different MDGA isoforms have functions that are distinct as well. First we needed to develop an assay that could be used to study and compare the function of both MDGAs. Since both MDGAs bind Nlgn2 in vitro and exert their functions via Nlgn2, it seemed reasonable to assume that this could serve as a starting point for developing this assay. Neuroligin 2 is a postsynaptic organizer of the inhibitory synapse (binding inhibitory post-synaptic proteins such as gephyrin) (Chubykin et al., 2007; Poulopoulos et al., 2009), and is a synaptogenic protein that in artificial synapse assays can recruit presynaptic terminals to cells that overexpress neuroligin 2 (Graf et al., 2004). In these assays, a transfected cell overexpressing the synaptogenic protein of interest is plated together with other neurons, resulting in the formation of “artificial synapses” formed at sites of contact between neurons and the overexpressing cell, visualized by an increase in markers of presynaptic terminals such as synapsin. This assay, first developed by Peter Schieffle, has been utilized by many other labs to identify adhesion
proteins that are synaptogenic and involved in synapse formation (Biederer and Scheiffele, 2007). Despite the fact that endogenous neuroligin 2 is localized to inhibitory synapses and binds to numerous post-synaptic inhibitory synapse proteins (Graf et al., 2004; Poulopoulos et al., 2009; Varoqueaux et al., 2004), when it is overexpressed in such an assay, the result is that both vglut1 and vgat positive presynaptic terminals are recruited to these artificial synapses (Linhoff et al., 2009; Takahashi et al., 2012), a result observed by multiple publications. When MDGA1 is coexpressed with neuroligin 2, it was found to suppress the presynaptic clustering of synapsin I presynaptic puncta around the MDGA1 and neuroligin 2 overexpressing cell, attenuating its synaptogenic activity (Lee et al., 2013; Pettem et al., 2013). Since the MDGAs function via NLG2, we decided to use neuroligin 2’s synaptogenicity as an assay for reading out MDGA2’s activity.

We first developed conditions for a synaptogenesis assay by overexpressing synaptogenic proteins such as Nlgn2 or Nlgn1, or non-synaptogenic surface proteins such as HRP-TM, together with a venus transfection marker (Fig. 4-4). In this assay, we chose to assay for vGlut1 and vGAT instead since we were eventually interested in how each MDGA might affect each of those synapse types specifically. Overexpression of Nlgn1 or 2 resulted in an increase in recruitment of presynaptic terminals containing vGAT or vGlut1, but not when HRP-TM or venus alone was expressed. This assay did not work well when the constructs were expressed under the synapsin promoter, which is a lot weaker for overexpression. Satisfied that the synaptogenesis assay worked, and that neuroligin 2 or 1 did not have a specific preference in the recruitment of a specific synapse type, we continued to apply this assay towards the MDGAs.
Figure. 4-4. Validation of synaptogenesis assay. Overexpression of either NLGN1 or NLGN2 leads to enhanced recruitment of both excitatory and inhibitory synaptic vesicles along transfected neurons. NLGN2, NLGN1 or HRP-TM was overexpressed in neurons (with whole cell Venus as a transfection marker). Enhanced recruitment of excitatory or inhibitory synaptic vesicles to transfected neurons was assessed by staining with anti-vGlut1 or anti-vGAT antibody, respectively. HRP-TM is used as a surface targeted non-synaptogenic negative control construct.

To explore and compare the role of both proteins in synapse formation, we performed a gain-of-function synaptogenesis assay by co-overexpression of Nlgn2 with or without Mdga1/2. We overexpressed these proteins in cortical neurons by lipofection under the CAG promoter, and assayed for the presynaptic glutamatergic and GABAergic markers vGlut1 and vGAT respectively. As shown in Fig. 4-5, overexpression of Nlgn2 by itself enhances the recruitment of both excitatory (vGlut1-type) and inhibitory (vGAT-type) neurotransmitter vesicles to opposing membranes, consistent with previous reports (Takahashi et al., 2012). Untransfected neurons do not recruit synaptic vesicles, suggesting this is a synaptogenic effect specific to overexpression of Nlg2. As expected, when Mdga1 was co-overexpressed with Nlgn2, the synaptogenic activity of Nlgn2 was suppressed, and the enhanced recruitment of presynapses containing both vesicle types was suppressed. Surprisingly, co-overexpression of Mdga2, however, suppressed selectively the recruitment of presynaptic terminals containing excitatory but not inhibitory vesicles (quantified in Fig. 4-8).
Figure. 4-5. Synaptogenesis assay based on overexpression of NLGN2 to probe specificity of presynaptic vesicle recruitment. NLGN2 was overexpressed in neurons alone (with whole cell Venus as a transfection marker) or together with Venus tagged-MDGA1 or MDGA2. Enhanced recruitment of excitatory or inhibitory synaptic vesicles to transfected neurons was assessed by staining with anti-vGlut1 or anti-vGAT antibody, respectively. Images are normalized to the same intensity scale and are representative of >20 transfected neurons per condition.

As a control we also assayed if MDGA2 might exert these effects by altering surface trafficking of neuroligin 2. This was tested by staining for the surface pool of neuroligin 2 (which is tagged with the V5 epitope) in a matched experiment setup with Fig. 4-6. No significant differences in the surface pools of neuroligin 2 were observed (images are normalized between conditions for each channel), both in the absence or when MDGA1/2 are coexpressed suggesting that MDGA expression does not affect neuroligin 2 surface levels, a result also previously observed for MDGA1 (Pettem et al., 2013).
Figure 4-6. Co-overexpression of MDGA1 or MDGA2 does not alter the surface levels of V5-NLGN2. Images are normalized and matched to samples shown in (E) and Fig. 4-5.

**Loss of function experiments to determine MDGAs effect on synapse density**

Based on the distinct activities for the two MDGA isoforms under co-expression with neuroligin 2, we hypothesized that Mdga1 and 2 both downregulate the trans-synaptic vesicle recruiting activity of Nlgn2 (which occurs via unknown presynaptic binding partner(s). However, Mdga1 binds Nlgn2 in such a way that it blocks recruitment of both inhibitory and excitatory vesicles. MDGA2 binds Nlgn2 differently, blocking recruitment of only excitatory vesicles while allowing recruitment of inhibitory vesicles. To test this hypothesis, we performed shRNA knockdown of Mdga1, Mdga2, or both together (Fig. 4-8) using shRNA hairpins previously described (Lee et al., 2013). Knockdown efficiency was quantified by qPCR in Figure 4-7. The interpretation of the knockdown data requires the assumption that each Nlgn2 binds to Mdga1 or 2, but not to both at once. In the case of Mdga1 knockdown, more Nlgn2 is freed to interact with Mdga2 instead. According to our model, Mdga2 promotes inhibitory vesicle recruitment but not excitatory. Correspondingly, we observe that Mdga1 knockdown causes an increase in inhibitory vesicle signal, but not excitatory
signal, in agreement with previous observations (Lee et al., 2013; Pettem et al., 2013). Mdga2 knockdown by itself has no significant effect, but when combined with Mdga1 knockdown causes both inhibitory and excitatory vesicle signals to increase. Hence, the double knockdown gives a phenotype distinct from either single knockdown, and results in full derepression of Nlgn2 activity.

![Graph](image)

**Figure.** 4-7 Relative mRNA abundance after KD of each MDGA by shRNA mediated RNA interference. Neurons were infected with lentiviruses driving expression of shRNA hairpins to MDGA1, MDGA2 or both MDGAs.
Figure 4-8. MDGA gain of function and loss of function assays. (A) Quantitation of coexpression of images in Figure 4-5 along with 8 additional fields of view per condition. Vesicle density is defined as the total anti-vGlut1 or anti-vGAT staining intensity divided by neuron area. Errors, ± s.e.m. *** indicates p < 0.0001 (student's t test). (B) Effect of single or double knockdown of MDGA1 and MDGA2 on excitatory and inhibitory vesicle density, quantified as in (A). Knockdowns were verified by qPCR in Fig. 4-7. 15 fields of view were analyzed per condition. Errors, ± s.e.m. ** indicates p < 0.01. (C) Effect of MDGA1 or MDGA2 overexpression (without NLGN2 co-overexpression) on excitatory and inhibitory vesicle densities, quantified as in (A). 12 fields of view were analyzed per condition. Errors, ± s.e.m. ** indicates p < 0.02.

In addition to the knockdowns, we also performed a gain-of-function assay by overexpressing only Mdga1 or Mdga2 (Fig. 4-8). More Mdga2 in neurons might shift the equilibrium for Nlgn2, causing more of it to bind to Mdga2 than Mdga1. Accordingly we would expect to see increased inhibitory vesicle recruitment, with no effect on excitatory vesicle recruitment. Our data in Fig. 4-8 shows this expected trend. Likewise, we found that MDGA1 overexpression did reduce the average vGlut1 and vGat density, suggesting a suppression of neuroligin 2's synaptogenecity, but we were also unable to see a significant difference in this experiment.

MDGAs domains responsible for its synaptic localization
Because our data demonstrate that MDGA1 and 2 have differing functions and localizations, we wanted to address which MDGA domains might be responsible for these effects. We generated a series of MDGA chimeras, swapped at the Ig 1-3 domains, Ig 1-6, at the Fibronectin III domain and also at the MAM and GPI anchor (Fig. 4-9). We tagged this panel of MDGA chimeras at the N terminus with HRP, to visualize their localizations at synapses using BxxP labeling. One of the chimeras, annotated B, did not traffic to the cell surface and was not considered further. We imaged the localization of these chimeras with respect to excitatory and inhibitory synaptic markers (Fig. 4-10), and compared these localizations to that of MDGA1 and MDGA2, asking which MDGA isoform were they more similar to. Quantification of this panel of Mdgal/2 chimeras (Fig. 4-11) showed that chimeras A, D, F were more MDGA2 like, while chimera C and E were more MDGA1 like. This suggest that their extracellular juxtamembrane Ig4-6 regions were responsible for their unique excitatory versus inhibitory synapse localizations (Fig. 4-9). Swapping the Ig1-3 domains, previously shown to be sufficient for binding Nlgn2 in cis (Lee et al., 2013; Pettem et al., 2013), did not alter their localizations, indicating that interactions with Nlgn2 do not determine the synapse subtype specificities of the Mdgas. This helps to reconcile the observation that Mdga1 is present in the excitatory cleft, while Nlgn2 is localized to inhibitory synapses (Levinson et al., 2005). Collectively, our experiments allowed us to identify Mdga2’s surprising and distinct function, and relate each Mdga’s role in regulation of inhibitory synapses to their respective localizations.
Figure. 4-9. Chimeras of MDGA1 and MDGA2 tested in (G). The parent genes each have 6 immunoglobulin (Ig) domains, a fibronectin type III (FNIII) domain, a memrin/A5 protein/receptor tyrosine phosphatase mu (MAM) domain, and a C-terminal GPI anchor. Chimera B exhibited poor surface trafficking and was not evaluated further.

Figure. 4-10. Imaging of the localization of Mdgal/Mdga2 chimeras. N-terminal HRP-tagged chimeras (see Fig. 4-9 for chimera description) were introduced by lentivirus into DIV15 rat cortical neurons. At DIV19, neurons were labeled live with BxxP, then fixed and stained with
neutravidin-AlexaFluor647, anti-vGlut1, and anti-VGAT to visualize the HRP fusion constructs, excitatory synapses, and inhibitory synapses, respectively. Scale bars, 10 μm. Quantitation of these data (along with additional 5 fields of view not shown, >300 puncta per construct) in Figure 4-11, and also displayed within images. Errors, ± 1 s.d.

![Bar chart showing colocalization](image)

**Figure. 4-11. Quantitation of the relative localization of chimeras to excitatory versus inhibitory synapses.** Assessed by imaging with anti-vGlut1 and anti-vGAT staining (images and error values shown in Fig. 4-10).

**Model for MDGA function and implications**

Based on these observations, we hypothesized that Mdga1 and 2 both downregulate the trans-synaptic vesicle recruiting activity of Nlgn2 (which occurs via unknown presynaptic binding partner(s) (Fig. 4-12)). However, Mdga1 binds Nlgn2 in such a way that it blocks recruitment of both inhibitory and excitatory vesicles. MDGA2 binds Nlgn2 differently, blocking recruitment of only excitatory vesicles while allowing recruitment of inhibitory vesicles. Mdga2 localization in the inhibitory synapse, serves to convert Nlgn2 into a specific recruiter of inhibitory synapses, to match the correct presynaptic content (GABA containing vesicles) to the correct inhibitory post-synaptic content (Nlgn2 is an organizer and recruiter of GABA<sub>A</sub> receptors, via its binding to gephyrin the inhibitory post-synapse scaffold protein (Poulopoulos et al., 2009)). Mdga1 in the inhibitory synapse serves to regulate inhibitory synapse size, much like the Nogo receptor complex in the excitatory synapse, while its role in the excitatory synapse may
serve as a "gatekeeper" to prevent recruitment of inhibitory post-synaptic elements via Nlgn2.

Figure. 4-12. Model for MDGA2 function at inhibitory synapses. Through unknown presynaptic binding partners, Nlgn2 can recruit both inhibitory and excitatory presynaptic terminals (left panel). We hypothesize that Mdga2 binds to Nlgn2 in *cis* to selectively block its recruitment of excitatory vesicles but not inhibitory vesicles (middle). In contrast, Mdga1 binds to Nlgn2 in *cis* to block recruitment of both vesicle types (right panel). This activity may serve to both regulate inhibitory synapse size and prevent invasion of Nlgn2 into excitatory synapses.

In support of this model, we have recently obtained preliminary data where we image the recruitment of vGlut1 at inhibitory synapses marked by gephyrin, and found that when Mdga2 is knocked out, there is a slight increase in vGlut1 recruitment to these synapses (Fig. 4-13).
Figure. 4-13. MDGA knockdown results in an increase in the accumulation of vGlut1 at Gephyrin sites. (A) Distribution of vGlut1/gephyrin intensity ratios, at gephryn sites. Endogenous gephryn was visualized using a mApple fusion to a gephryn intrabody (Gross et al., 2013). shRNAs against MDGA1, 2, or both, and gephryn-intrabody were introduced by lentiviral transduction at DIV7. At DIV15, rat cortical neurons were fixed and stained with anti-vGlut1 and anti-vGAT antibody. Distributions were compared using Tukey’s HSD test. (B) Representative images of gephryn visualized by the mApple intrabody (red), within shRNA expressing neurons co-expressing GFP, together with anti-vGlut1 (blue) or anti-vGAT(green) staining.

Conclusions

We have used our synaptic cleft proteomes (discussed in Chapter 3) to identify protein isoforms from the same family (and hence likely similar function) that have a unique segregation to excitatory and inhibitory synapses. The observation of Mdga1 and Mdga2, previously assumed to be identical in
localization and function (Lee et al., 2013; Pettem et al., 2013), in our excitatory and inhibitory proteomes, respectively, inspired us to investigate Mdga2, which had not been studied in neurons before. We were able to formulate and test the hypothesis that Mdga2, discovered in our inhibitory cleft proteome, plays a novel and unexpected role in inhibitory synapse formation. Overexpression and knockdown experiments gave results consistent with a model (Fig. 4-12) in which Mdga1 and 2 compete for cis binding to Nlgn2 on the post-synaptic membrane. When Mdga1 is bound to Nlgn2, the latter is inhibited from using trans-synaptic mechanisms to recruit either inhibitory vesicles or excitatory vesicles. When Mdga2 is bound to Nlgn2, only excitatory vesicle recruitment is blocked. Hence Mdga2 appears to be a specificity factor, conferring to Nlgn2 the ability to template specific construction of inhibitory synapses with proper matching between pre-synaptic termini containing GABA-type inhibitory vesicles and GABAergic post-synaptic termini. Hence Mdga2 is one part of an answer to the search for molecular mechanisms that control selective pre-to-post synaptic matching in the process of inhibitory synapse formation. Mdga1, localized to the excitatory cleft and possibly inhibitory cleft as well, has a different role: to attenuate the activity of Nlgn2 overall, perhaps to prevent invasion of inhibitory synapse elements into excitatory synapses. These results illustrate how we can use spatially-resolved live-cell proteomic mapping to help address longstanding molecular questions in neuroscience.

**Materials and Methods**

**NLGN2 synaptogenesis assay**

Cortical neurons plated on coverslips in 24 well plates as described under “Rat cortical neuron culture”, were transfected at DIV 12 using 1ul of Lipofectamine 2000 per well, with 100 ng of pCAG-V5-NLGN2A plasmid DNA, together with 100 ng of either pCAG-venus-MDGA1, pCAG-venus-MDGA2 plasmids, or on its own
for 3 hours. The media was then replaced with neuron culture media (Neurobasal containing B27, Glutamax, penicillin, and streptomycin and FUDR) and cells were cultured for another 2 days before fixing and permeabilized for immunostaining with mouse-anti-vGlut1 and rabbit anti-vGAT. For Fig. 4-5 neurons were processed similarly but transfected with 10 ng of pCAG-venus plasmid, with either no additional DNA, or 100 ng of either pCAG-V5-NLGN2A, pCAG-V5-NLGN1 or pCAG-HRP-TM plasmid DNA for 3 hours instead.

RNA interference knockdown of MDGA1 and MDGA2

For RNA interference knockdown by plasmid-based short-hairpin RNA (sh-RNA), the oligonucleotides that target nucleotides 1027–1045 of rat MDGA1 (5′-GTCTCTTTCTTCTACCACA-3′) or 1933–1951 of rat MDGA2 (5′-AGGTGAAGCTAAAGAACAA-3′), were subcloned into LenLox3.7 variant pLLs-GFP to express GFP and sh-MDGA1 under the human synapsin promoter and U6 promoter, respectively. These hairpins had previously been used to knockdown MDGAs (Lee et al., 2013). We used sh-MORB (5′-GATGGTGGCAGTACCAGTG-3′) as a control sh-RNA (sh-scrambled), which has no effects on neuronal morphology (Pettem et al., 2013). Lentiviruses were generated as described under “Generation of lentiviruses for trans-gene expression in cortical neurons and neuron infection protocol”.

Cortical neurons were cultured as described under “Rat cortical neuron culture”. Cortical neurons at days in vitro DIV 5 were infected with lentiviruses and the samples were fixed at DIV 19 with 4% paraformaldehyde in “fixation buffer” at room temperature for 10 minutes. The cells were then washed three times with DPBS.

Cells were permeabilized with 0.1% TritonX-100 in DPBS for 7 minutes at room temperature. After three more washes with DPBS, the cells were blocked with 3% w/v bovine serum albumin (BSA) in DPBS for one hour, then stained with
mouse anti-vGlut1 (NeuroMab, 1:1000) and rabbit anti-vGAT (Synaptic Systems, 1:1000), in 3% w/v BSA in DPBS for 2 hour at room temperature or overnight at 4 °C. Samples were then washed 3 x 1 minutes with DPBS. Samples were then incubated with secondary goat anti-mouse-Alexa-Fluor-647 (AF647) and goat anti-rabbit-Alexa-Fluor-568 antibody (AF568) (Invitrogen, 1:1000 dilution each) in 3% w/v BSA in DPBS for one hour, then washed 3 x 1 minutes with DPBS. Samples were then imaged by confocal microscopy with 5 x 0.3um stacks using a 100x objective.

**MDGA overexpression assay**

Cortical neurons plated on coverslips in 24 well plates as described under “Rat cortical neuron culture”, were transfected at DIV 12 using 1ul of Lipofectamine 2000 per well, with 50 ng of either pCAG-venus-MDGA1, pCAG-venus-MDGA2 plasmids, or a venus transfection marker (as a control) for 3 hours. The media was then replaced with neuron culture media (Neurobasal containing B27, Glutamax, penicillin, and streptomycin and FUDR) and cells were cultured for another 4 days before fixing and permeabilized for immunostaining with mouse-anti-vGlut1 and rabbit anti-vGAT.

**Imaging NLGN2 surface expression**

To detect surface expression of V5-tagged NLGN2A, fixed, unpermeabilized cells were blocked at room temperature with 3% w/v BSA in DPBS for one hour at room, then stained with mouse anti-V5 (Invitrogen, 1:1000 dilution) for one hour, then washed three times with DPBS. Samples were then incubated with secondary goat anti-mouse-Alexa-Fluor-647 (AF647) antibody (Invitrogen, 1:1000 dilution) in 3% w/v BSA in DPBS for one hour at room temperature, then washed 3 x 1 minutes with DPBS. Samples were then imaged by confocal microscopy using a 40x objective.
Image analysis
To analyze the synaptic density a mask for the venus channel was crossed with either a mask for the vGlut1 or vGAT channel to generate new masks of their overlapping regions. The fluorescence intensity of the vGlut1 or vGAT channels that overlapped with the venus channel was averaged over the area of the entire venus mask, giving the average intensity of the synaptic marker recruited over the transfected neuron. This was then normalized to the control condition to obtain the relative densities.

References


Varoqueaux, F., Jamain, S., and Brose, N. (2004). Neuroligin 2 is exclusively

Chapter 5: Mapping the Axon initial segment using APEX

Sections of this Chapter were carried out in collaboration with an MIT undergraduate Kayvon Pedram.
Introduction

The directionality of signal transduction within a neuronal circuit is in part conferred by the polarity of its individual components; neurons are highly polarized cells, with signals arriving at post-synapses formed along the somatodendritic processes, and these signals are transferred along the axon into presynapses that connect that neuron to the next signal receiving cell. The axon initial segment (AIS) is an important subcellular structure within each neuron (Fig. 5-1) that is intrinsically established (without external cues) that demarcates a region that performs the highly specialized function of integrating all the signals that a neuron has received from its post-synapses, which in turn determines if a signal is passed along the circuit to its receiving cells via an action potential, which also begins within the AIS (Rasband, 2010). This 20-40μm structure is assumed to contain a unique set of machinery that is involved in carrying out its integrative function (including voltage gated ion channels that initiate the action potential, scaffolding molecules and other signaling components) and has been shown to contain both membrane (Winckler et al., 1999) and cytoplasmic diffusion barriers (Song et al., 2009) because it delineates the end of the signal receiving somatodendritic region of the cell and the beginning of the signal sending axon. Because the AIS is a relatively large structure (on the orders of 1-2 microns) it is well resolved by confocal microscopy and is as important a subcellular structure as the synaptic compartments (cleft, active zone, PSD). This structure is completely impossible to purify by biochemical means nor approximate since it is not membrane bound and does not possess any unique biochemical properties that can be used to isolate it.
Figure 5-1. The axon initial segment (AIS). The AIS is located at the most proximal region of the axon as indicated by the box region, functioning as both a physical and physiological bridge between the somatodendritic and axonal regions of the neuron.

We thought that we would be able to leverage the spatial resolution of the APEX labeling technology (Hung et al., 2014; Rhee et al., 2013) using the cell permeable probe biotin-phenol (BP), towards identifying the proteins components of this extremely interesting region as shown in the scheme below (Fig. 5-2). Questions of what structural components comprise the cytoplasmic and membrane filter, how does the structure establish itself intrinsically, what protein components comprise the structure remain unanswered, despite the importance of this extremely interesting subcellular structure unique to neurons.

![Scheme for mapping the Axon initial segment without organelle purification.]

APEX is genetically targeted to the AIS. The grey shapes are endogenous proteins residing inside and outside the AIS. To initiate labeling, a BP is added to the live neurons for 30 minutes to allow sufficient penetration into the cell, and H$_2$O$_2$ is added for 1 minute to initiate labeling. The enzyme catalyzes the formation of reactive biotin species, which covalently tag proximal endogenous AIS proteins. Subsequently, neurons are lysed and biotinylated proteins are isolated using streptavidin (SA) beads for identification by mass spectrometry (MS).

Targeting APEX to the Axon initial segment

To target APEX to the AIS, there were two strategies that could be explored; we could fuse APEX directly to proteins at the AIS, as had been done
for the synaptic cleft, or identify protein domains that are responsible for localizing native AIS proteins to the AIS. A major component of the AIS is the structural protein ankyrinG (Jenkins and Bennett, 2001; Zhou et al., 1998), of which many AIS proteins such as the cell-adhesion protein neurofascin NF-186 bind via its C terminal FIGQY domain (Garver et al., 1997). The proteins responsible for initiating an action potential at the AIS are the voltage gated sodium channels (VGSCs) (Kole et al., 2008). The targeting of these channels to the AIS has been well studied and is known to occur by via an ankyrinG binding motif that contains a critical 9 amino acid motif shared by the sodium voltage gated ion channels Nav 1.1, 1.2, 1.3 and 1.6 (Garrido et al., 2003). This motif is also homologously shared with KCNQ2/3-type potassium channel subunits that also target to the AIS via binding to ankyrinG (Pan et al., 2006; Rasmussen et al., 2007). Also crucial to accurate targeting is a removal sequence, that is found within the loops 2-3 of the Navs, that targets these channels for selective endocytic removal (Fache et al., 2004), from membrane sites not corresponding to the AIS, and also kinase CK2 phosphorylation sites that regulate the binding to ankyrinG (Bréchet et al., 2008). Because these components were the minimum requirements needed for proper AIS targeting, and highly conserved between many different Nav subunits, it is not hard to imagine that a chimeric protein fused to this loop might be sufficient for efficient targeting to the AIS; a fusion of this loop with the extracellular domain of CD4 (Garrido et al., 2003), extracellular domain of neurofascin (Lemaillet et al., 2003), GFP (Garrido et al., 2003) or ChR2 (channel rhodopsin 2, (Grubb and Burrone, 2010)) already targets these proteins efficiently to the AIS.

We decided to fuse APEX to the loops 2-3 of the NaV voltage gated ion channel 1.2 (abbreviated NaVII-III) to target APEX to the AIS. This was cloned into a lentiviral vector with expression under the synapsin promoter (see Chapter 2). This turned out to be a fortuitous choice, since even though under relatively high expression levels (by lipofection) it was possible to observe some occasional, targeting to the AIS in cortical neurons but most surprisingly; the biotinylation pattern as generated when biotin-phenol and H₂O₂ was added was
very tightly localized to what appeared to be the AIS (Fig.5-3). This encouraged us to progress towards generating lentiviruses for the efficient targeting of many neurons needed for proteomics, and also to access a lower expression level regime needed for proper targeting.

Figure 5-3. APEX-Navil-Ill lipofected into DIV 18 cortical neurons gives a biotinylation signal that occasionally, is tightly localized to the AIS. Labeling was carried out using 500μM Biotin phenol, preincubated for 30 min in Tyrodes buffer followed by 1mM H₂O₂ addition for 1 min, then fixed, permeabilized and immunostained for V5 and biotinylation using neutravidin 647 conjugate (Nv-647).

**Lentiviral infection was effective at reducing the non-specific localization of APEX, and at this point, a former graduate student Dr. Stephanie Lam had evolved a more active variant of APEX, APEX2 (Lam et al., 2014), which we used instead of the original APEX for the subsequent experiments. To validate that our APEX2- NaⅡ-Ⅲ construct was targeted to the AIS and could label the structure, we carried out co-immunostaining with endogenous protein markers to demonstrate that 1) the biotinylation pattern was excluded from the somatodendrites of neurons, by using an antibody to the somatodendritic marker MAP2, which is also excluded from the neuronal axon, and 2) that the labeling itself corresponded with a marker of the AIS, the structural protein ankyrinG. We**
were able to show that the labeling was well excluded from the soma and dendrites (Fig. 5-5), and corresponded cleanly with the axon initial segment, as visualized by ankyrinG (Fig. 5-4). The labeling also did not extend throughout the axon, suggesting that the labeling radius itself was short enough to label the AIS specifically. The labeling itself was also sensitive to expression levels; when lipofection is used many cells tended to overexpress and mislocalize the construct/labeling (Fig. 5-3), and only when lentiviruses are used does the labeling become consistently tight (see Fig. 5-4 and 5-5).

Figure 5-4. Lentiviral infected DIV 18 cortical neurons expressing APEX2-Nav\textsubscript{II}-III give a biotinylation signal that looks tightly localized to the AIS and colocalizes well with AnkG. Cells were infected at DIV 14 and labeled at DIV 18. Labeling conditions are the same as Fig. 5-3, and neurons were additionally stained for endogenous Ankyrin G protein.

Figure 5-5. BP labeling of the Axon initial segment by APEX2-Nav\textsubscript{II}-III is excluded from somatodendritic processes as shown by exclusivity with somatodendritic marker MAP2. Cells were infected at DIV 14 and labeled at DIV 18. Labeling conditions are same as Fig. 5-3, and neurons were additionally stained for endogenous MAP2 protein.

APEX2-Nav\textsubscript{II}-III gives a distinct biotinylation fingerprint of proteins at the AIS

In preparing for a proteomics experiment, we verified that the AIS targeted
APEX2 biotinylates a unique set of proteins by comparing the streptavidin western blot fingerprint generated by APEX2-Nav1.2I-III with APEX/HRP targeted to distinct subcellular compartments in the cell (Fig. 5-6). We followed up by scaling to a larger scale in terms of number of neurons used (1 10cm dish, 3 million neurons per condition) and enriched for proteins biotinylated by APEX using streptavidin magnetic beads and these beads were washed under the conditions used in Rhee et al. for labeling the mitochondrial matrix. We were able to enrich for proteins hopefully at the AIS, although the banding patterns only looked darker in intensity by coomaissie and silver staining than the negative controls, and there were not distinctly obvious differences by the NES fusion’s labeling (Fig. 5-6).

What seemed promising however was that the protein MAP2 was enriched to a smaller extent by APEX2-Nav1.2I-III, than by APEX2-NES (Fig. 5-7). MAP2 is a protein excluded from the AIS and should be labeled by the NES construct but not by APEX2-Nav1.2I-III. This suggested that we were labeling and enriching the different subcompartments (AIS versus the general cytosol) in a spatially defined manner. Since fluorescence microscopy imaging the biotinylation pattern was specific to the AIS, and proteins were sufficiently biotinylated to generate enough material for proteomics using the pulldown conditions developed, we decided to perform a mass spectrometry based proteomics experiment.
Figure 5-6. Gel analysis of streptavidin-enriched lysates after live-neuron biotinylation with APEX2-Na
II-III, visualized by silver staining and coomaissie. Labeling for 1 minute with BP was performed as in Fig. 5-3. Cell lysis, streptavidin enrichment, and elution from beads was performed as in (Rhee et al., 2013). Eluate was run on 8% SDS PAGE and visualized by silver stain (left) or coomaissie (right).

Figure 5-7. Streptavidin blot (left) and MAP2 (right) enrichment analysis of streptavidin-enriched lysates after live-neuron biotinylation with APEX2-Na
II-III. MAP2 western blot shows lanes with either whole cell lysate (prior to streptavidin bead enrichment) or eluate from streptavidin bead enrichment for each of the APEX fusions/negative control. This blot shows a differential enrichment of MAP2 protein, a protein not found at the AIS.
Mapping the proteome of the Axon initial segment

Quantitative proteomics for discovering the axon initial segment proteome

We decided to approach the AIS proteome using two different quantitative proteomics methods (note that these experiments occurred before the synaptic cleft proteomics experiments described in Chapters 2-3, and were used to gain insights into developing conditions for the eventual synaptic cleft proteomes). The two most well established methods at this point were iTRAQ and SILAC, iTRAQ being more preferable for labeling primary cells that cannot metabolically incorporate the isotopically labeled amino acids into their proteome, while SILAC (Ong et al., 2002) has the advantage of being less sensitive to sample handling since the key quantitative mixing step occurs early on prior to enrichment (See Chapter 2 for a detailed discussion of the use of iTRAQ for proteomics). Our workflow for the experiments using each method is shown in Fig. 5-8 and 5-9.

Briefly, we performed 3 quantitative proteomics experiments per quantitative proteomics methodology; SILAC or iTRAQ (6 experiments in total), and each of the 3 had the following paired conditions (Fig. 5-8 and Fig. 5-9): the first pair were APEX targeted to the AIS compared to a negative control where APEX is omitted, the second pair is APEX targeted to the AIS, compared to a negative labeling control where the labeling substrate BP is omitted, and the last pair was APEX targeted to the AIS compared to APEX targeted to the cytosol using a nuclear excluded sequence peptide fused to APEX (APEX2-NES).
Characteristics and analysis of the proteomic data

Although the imaging data controls that were performed in parallel to our experiment looked very promising, we found that the actual mass spec data revealed a number of flaws in the dataset, which eventually resulted in us being unable to use this to define an AIS "proteome", these flaws however encouraged us to improve the APEX methodology, and were applied towards the final synaptic cleft proteomes discussed in Chapter 2-3.

Each iTRAQ and SILAC experiment revealed the following numbers of
detected proteins as shown in Fig. 5-10. We first removed any proteins that did not have 2 or greater detected unique peptides (a standard in proteomics for genuine detection), and then renormalized the iTRAQ or SILAC ratios. We plotted the distributions for each experiment as a function of their ratios in a histogram, and found that they appeared mostly normal distributed, instead (Fig. 5-11) of having a shoulder on one end, like in the IMS or mitochondrial matrix proteomes (the distribution shown belongs to the SILAC experiment). This could be explained by a number of explanations, one that we did not enrich for biotinylated proteins successfully or the presence of many non specific bead binders during the enrichment, or alternatively that the AIS proteome was of such low complexity (the absolute number of unique proteins within the AIS is low), that the non-specifically bound proteins would represent the majority of the distribution, with a small pool of genuine AIS proteins representing the right enriched tail of the distribution (Fig. 5-11).

Figure 5-10. Number of proteins identified in each quantitative proteomics experiment based on different unique peptide cutoffs.
Figure 5-11. Distribution of proteins in experiment (SILAC experiment is shown) that resemble a normal distribution with a slight right leaning shoulder.

The correlation between experiments was not high, and $R^2$ values for the correlation curves between experiments 1 and 2 within each quantitative proteomics experiment were 0.66 for iTRAQ and 0.27 for SILAC (Fig. 5-12); this represents the paired experiment of APEX targeted to the AIS compared to its negative labeling control. This suggests there is poor reproducibility between the experiments, and which could be contributed by background contaminants, which are not biotinylated and hence not affected by the enrichment process, and hence have randomly generated and not correlative enrichment ratios, thus contributing towards a poor $R^2$ value.
We analyzed the data by first identifying proteins that were known residents of the AIS (AIS Gold+ list) and proteins that could not be labeled by AIS targeted APEX (AIS Gold- list contains proteins such as mitochondrial matrix proteins). We plotted a chart in R (Fig. 5-13), with log2 normalized SILAC ratios in descending order for all proteins detected with 2 or more unique peptides in all three experiment replicates (837 proteins total), that also had a SILAC ratio >0 for the third experiment (which involves the ratiometric tagging comparison) and this is presented on the left in Fig. 5-13, rationalizing this at least represents proteins that had been enriched at the AIS compared to the cytosol. Each protein is represented by one line, and shown in green, red or gray. Green lines represent “Gold+” proteins, which are proteins that we believe based on prior literature have residency at the AIS. Red lines represent “Gold-” proteins, which are proteins that cannot be present at the AIS because they cannot be biotinylated by AIS targeted APEX, (proteins such as those with strong nuclear or mitochondrial matrix Gene ontology annotation). Gray proteins are neither of the above; these proteins are the ones we hope to use to gain biological insight into the AIS.

From the chart, we found that proteins that were known to have residency within the AIS such as AnkG and neurofascin, and quite a number of voltage gated sodium channels and potassium channels were all enriched within the
proteomic list with good enrichment ratios (green bars in the R-chart plot in Fig. 5-13). However, even though we did find many of them to be enriched, not all known residents of the AIS proteins had high enrichment ratios, and would indicate that our coverage (as measured by how well we could enrich these known residents) was poor. We suspect that one explanation for this is proteins that are dual localized to the cytoplasm and the AIS are both labeled by our APEX fused baits, resulting in their lack of enrichment. One other explanation might be related to the choice of differential labeling bait (APEX2-NES), which might not be the most appropriate because it is not completely excluded from the AIS and can also label AIS proteins. Perhaps APEX fused to a somatodendritic marker MAP2 (or the smaller MAP2c) would be a better choice of bait, since it is occluded from the AIS.

Additionally we found that a number of members of the Gold+ list were not detected at all (greater than 2 unique peptides) within the proteomic experiment, suggesting the experiment’s coverage is poor in terms of being sensitive enough to detect the protein itself. This suggests that a larger input material would improve the quality of the proteomic data and we incorporated this into subsequent proteomic experiments (synaptic cleft proteomics in Chapter 2-3).

We also found that there were a significant number of proteins that could not be found in the AIS that were enriched in our dataset, and these included histone and mitochondrial matrix proteins. Although the bulk of these gold-proteins had enrichment ratios centered around 0, there were still a reasonable number that were enriched within the dataset, suggesting that the enrichment was not perfectly clean, and more validation for verifying a protein is at the AIS would be necessary downstream.

To improve the specificity of the resulting dataset, we applied cutoffs for SILAC ratios that were greater than 0, for all 3 experiments. In this way, we observe that many proteins that are non-specific binders are eliminated, and we can observe a clearer enrichment of AIS proteins (Fig. 5-13 middle panel) and
deenrichment of proteins that should not be in the AIS (gold- proteins). This is suggestive that the labeling and enrichment for the proteomics experiment did work to some extent. A more stringent cutoff is applied in the third panel on the right. These cutoffs however, also resulted in a loss of coverage (number of gold+ proteins retained) that is a tradeoff for improving the specificity of the dataset.

Our main concerns with regards to this list however are the following: clearly based on the histograms it seems that enrichment of biotinylated proteins is not occurring efficiently, and we are suffering from a high non specific background. There is also poor correlation between the repetition experiments, and based on two criteria, specificity (number of gold- proteins that are detected/total) and depth of coverage (number of gold+ list proteins detected/total), the dataset itself is not very impactful list, even when the specificity was high (using high ratio cutoffs) and depth of coverage was low, since many of the proteins present had no clear link to the AIS, and would be difficult to pursue a physiological function related to the AIS.

We speculate that some of the following reasons might explain some of the flaws observed in the dataset: 1) The AIS proteome may be very small in terms of absolute numbers of unique proteins, meaning that our true hits might be lost in the noise from nonspecific binders or abundant proteins that are only slightly biotinylated. 2) We have APEX- and H2O2-dependent BP labeling that is diffuse and APEX targeted to the AIS is also labeling proteins in the cytosol slightly during the 1 minute H2O2 incubation period. 3) We have APEX-dependent but H2O2-independent background labeling: the enzyme is labeling proteins during the 30 minute BP incubation period, perhaps with endogenous H2O2 present within the cell. This has the potential to spread biotinylation beyond the AIS. 4) Our SA enrichment is poor, resulting in unacceptable amounts on nonspecific binding. Since there were a number of non specific background proteins that were still enriched in our dataset, we decided to reinspect the pulldown steps in neurons, and optimize a protocol for obtaining good enrichment
and low background binding.

Figure 5-13. Plots showing the relative enrichments of various known true positive AIS proteins (shown in green) and the known true negatives, proteins that cannot be labeled by AIS targeted APEX (shown in red), while proteins in neither group are in grey, within the entire dataset. These charts were generated after applying cutoffs to the iTRAQ ratios for each experimental replicate: left chart (>0 for Exp03), middle chart (>0,0,0) or right chart (>0.1,0.1,0.25), where the numbers represent (Exp01 cutoff, Exp02 cutoff, Exp03 cutoff).

Optimizations for subsequent proteomic experiments

Optimization of enrichment of biotinylated proteins

To improve the enrichment of biotinylated proteins, we decided to tackle this from two avenues, 1) bead washes and 2) streptavidin bead to biotinylated input material ratios. To evaluate the efficiency of the wash conditions, we used densitometry to compare the amount of non-specific sticking to the streptavidin beads, identifying the harsher wash condition (lane 6 of Fig. 5-14) as the most
efficient at reducing the non specific sticking background, this bead washing condition was applied to subsequent proteomes (discussed in Chapter 2-3).

Untransfected

<table>
<thead>
<tr>
<th>ID</th>
<th>IDV</th>
<th>%</th>
<th>AREA</th>
<th>AVG</th>
<th>BACK</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>245941854</td>
<td>19.3</td>
<td>26727</td>
<td>9202</td>
<td>18541</td>
</tr>
<tr>
<td>2</td>
<td>230172924</td>
<td>18.1</td>
<td>26727</td>
<td>8612</td>
<td>18541</td>
</tr>
<tr>
<td>3</td>
<td>265345656</td>
<td>19.3</td>
<td>26727</td>
<td>9202</td>
<td>18541</td>
</tr>
<tr>
<td>4</td>
<td>209299137</td>
<td>20.9</td>
<td>26727</td>
<td>8612</td>
<td>18541</td>
</tr>
<tr>
<td>5</td>
<td>178188909</td>
<td>14.0</td>
<td>26727</td>
<td>6667</td>
<td>18541</td>
</tr>
<tr>
<td>6</td>
<td>142241094</td>
<td>11.2</td>
<td>26727</td>
<td>5322</td>
<td>18541</td>
</tr>
</tbody>
</table>

Figure 5-14. Using densitometry for optimization of wash conditions for streptavidin bead pulldown. Uninfected HEK cells were labeled with biotin-phenol as per normal, then lysed and applied to streptavidin beads for pulldown. The following conditions were used: 1) 2M urea (current protocol) 2) 4M urea (higher Urea wash) 3) 1M KCl (high salt wash) 4) 0.1M Na2CO3 pH 11.5 (high pH wash) 5) 1M KCl 0.1M Na2CO3 (high salt high pH) 6) 1M KCl 0.1M Na2CO3, 2M Urea (high salt high pH and urea). Gels were visualized by silver staining after pulldown.

We hypothesize that by titrating the beads to biotinylated input material ratio, we could minimize the amount of non-specific bead binding, which is directly correlated with the amount of streptavidin magnetic beads used. Since neurons are primary cells and do not divide, the number of AIS available for biotinylation and hence enrichment is directly proportional to the density of cells plated (each neuron typically has 1 AIS). By titrating the number of beads to neurons plated, we were able to identify conditions which resulted in good enrichment of biotinylated proteins, with minimal background, as observed by the enriched protein pattern (visualized by silver staining) resembling the biotinylated protein pattern (as visualized by streptavidin western blotting) as shown in Fig. 5-15).
Conclusions

In this chapter, we have developed reagents that could enable the biotinylation and isolation of proteins within the Axon initial segment by a proteomics experiment. Our first attempt at applying these reagents for proteomics was unfortunately plagued by a large number of false positives, and did not give good coverage of known AIS proteins. These results inspired us to improve the enrichment protocol for biotinylated proteins, and helped develop new wash conditions that were subsequently applied to other proteomes (Chapters 2-3) Using these new wash conditions, and a titration of biotinylated material to the minimum number of beads required for enrichment we revisited the enrichment of biotinylated AIS proteins and found that we were able to obtain conditions where the streptavidin western blot resembled the proteins eluted after enrichment and visualized by silver staining, suggesting that most of the proteins that we were eluting were actually biotinylated proteins. We are hopeful that this could eventually lead to a successful proteomic mapping of the axon initial segment.
Materials and Methods

Localization and activity of peroxidase fusions by fluorescence microscopy

Cortical neurons were cultured as described under “Rat cortical neuron culture”. Cortical neurons at days in vitro (DIV) 14 were infected with lentiviruses expressing the APEX-tagged AIS constructs at a titer previously empirically optimized (see “Titrating lentiviral constructs” in Chapter 2) to give localized BP labeling of the AIS.

For BP labeling, cells at DIV 18 were first incubated with 500 μM Biotin-phenol (BP) in Tyrode’s buffers for half an hour, followed by an addition of 1 mM H₂O₂ for 1 min. Samples were then quenched and fixed with 4% PFA as described in Chapter 2.

Cells were next permeabilized with 0.1% TritonX-100 in DPBS for 7 minutes at room temperature. After three more washes with DPBS, the cells were blocked with 3% w/v bovine serum albumin (BSA) in DPBS for one hour, then stained with mouse anti-V5 (Invitrogen, 1:1000 dilution), Rabbit anti-MAP2 (Abcam (ab32454), 1:1000), or Mouse anti-AnkG (clone 4G3F8 1:500 dilution). Samples were then washed 3 x 1 minutes with DPBS. Samples were incubated with neutravidin-AF647 conjugates (see Chapter 2), secondary goat anti-mouse-Alexa-Fluor-488 (AF488) or goat anti-rabbit-Alexa-Fluor-568 antibody (AF568) (Invitrogen, 1:1000 dilution each) in 3% w/v BSA in DPBS for one hour, then washed 3 x 1 minutes with DPBS. Samples were then imaged by confocal microscopy using a 40x objective.

Generation of biotinylated cell lysate for SDS-PAGE analysis

Two 10 cm dishes were prepared with 2.5 million cortical neurons in each. At DIV 14, neurons were infected with lentiviruses expressing the APEX-tagged AIS
construct. At DIV 18, first incubated with 500 μM Biotin-phenol (BP) in Tyrode’s buffers for half an hour, followed by an addition of 1 mM H₂O₂ for 1 min. Samples were then quenched with 3 washes of Tyrode’s buffer containing 10 mM sodium azide, 10 mM sodium ascorbate, and 5 mM Trolox. After quenching, the cells were harvested by scraping and pelleted by centrifugation at 3,000 g for 10 minutes. The supernatant was discarded and the pellet was stored at −80 °C overnight.

The cell pellet was lysed by incubating in 500μl of RIPA lysis buffer (50 mM Tris-HCl, pH 8.0, 150 mM NaCl, 0.2% SDS, 0.5% sodium deoxycholate, 1% Triton-X 100) (Hung et al., 2016; Rhee et al., 2013) containing protease inhibitor cocktail (Sigma Aldrich, catalog no. P8849), 1 mM PMSF (phenylmethylsulfonyl fluoride), 10 mM sodium azide, 10 mM sodium ascorbate, and 5 mM Trolox. Lysates were centrifuged at 16,000 g for 10 minutes at 4 °C, with the supernatant used for subsequent processing.

Streptavidin blotting of biotintylated neuron lysate

Lysates were combined with SDS protein loading buffer and boiled for 5 min, then run on a 8% SDS-PAGE gel. The gel was transferred to a nitrocellulose membrane. The membrane was stained with Ponceau S (10 minutes in 0.1% w/v Ponceau S in 5% (v/v) acetic acid/water) and imaged. After destaining in deionized water, the membrane was blocked with 3% w/v BSA in 1 x TBST (0.1% Tween-20 in Tris-buffered saline) at 4 °C overnight. The membrane was rocked in 0.3 μg/ml streptavidin-HRP (Thermo Scientific) in 1% w/v BSA in 1 x TBST at room temperature for 1 hour, then washed with 1 x TBST four times for 5 minutes each time. Finally, the blot was developed with Clarity Western ECL Substrate (Bio-Rad) and imaged on an Alpha Innotech gel imaging system.

Streptavidin enrichment of biotinylated lysate and analysis by silver staining
To enrich biotinylated proteins, streptavidin-coated magnetic beads (Pierce catalog no. 88817) were prepared by washing twice with RIPA lysis buffer. Then, 500 μl of whole cell lysate was incubated with 150 μl of streptavidin bead slurry overnight at 4 °C with gentle rotation. The beads were washed with 2 × 1 ml RIPA lysis buffer, 1 × 1 ml of 2 M urea in 10 mM Tris-HCl, pH 8.0, and again with 2 × 1 ml RIPA lysis buffer.

Biotinylated proteins were eluted by boiling the beads for 10 min in 50 μl 3× protein loading buffer supplemented with 20 mM dithiothreitol (DTT) and 2 mM biotin. The streptavidin eluate (SAE) was collected and run on a 8% SDS-PAGE gel. To analyze biotinylated proteins, western blots were performed as described above. To analyze all protein material eluted, gels were silver stained (Pierce) instead, then imaged on an Alpha Innotech gel imaging system.

Other chemicals and reagents

BP labeling and reagents are described in reference (Hung et al., 2016; Rhee et al., 2013). All chemicals were purchased from Sigma-Aldrich unless otherwise specified.

References


Proteomics 1, 376–386.
Chapter 6: Enzymatic probe ligation of neurexin-neuroligin interactions in neurons

Introduction

Neurexin and neuroligin are transmembrane adhesion proteins that play an important role in organizing the neuronal synaptic cleft. Our lab previously reported a method for imaging the trans-synaptic binding of neurexin and neuroligin called BLINC (Biotin Labeling of INtercellular Contacts). In BLINC, biotin ligase (BirA) is fused to one protein while its 15-amino acid acceptor peptide substrate (AP) is fused to the binding partner. When the two fusion proteins interact across cellular junctions, BirA catalyzes the site-specific biotinylation of AP, which can be read out by staining with streptavidin-fluorophore conjugates. We discovered that BLINC in neurons cannot be reproduced using the reporter constructs and labeling protocol previously described. We explored the technical reasons for the lack of reproducibility and then re-design the BLINC reporters and labeling protocol to achieve successful neurexin-neuroligin BLINC imaging in neuron cultures. In addition, we introduce a new method, based on lipoic acid ligase instead of biotin ligase, to image trans-cellular neurexin-neuroligin interactions in human embryonic kidney cells and in neuron cultures. This method, called ID-PRIME for Interaction-Dependent PRobe Incorporation Mediated by Enzymes, is more robust than BLINC due to higher surface expression of lipoic acid ligase fusion constructs, gives stronger and more localized labeling, and is more versatile than BLINC in terms of signal readout. ID-PRIME expands the toolkit of methods available to study trans-cellular protein-protein interactions in living systems.

Methodologies for labeling trans-synaptic interactions

Neurexins (NRX) are presynaptic adhesion proteins that bind across the synaptic cleft to postsynaptic neuroligins (NLG). This trans-cellular binding is believed to play a role in synapse formation, specification, and/or stabilization (Craig and Kang, 2007). To facilitate the study of NRX-NLG biology, it would be desirable to have a non-invasive method that reports on their binding in living
cells. A recent study (Yamagata and Sanes, 2012), building upon the GRASP technology (GFP Reconstitution Across Synaptic Partners) introduced earlier (Feinberg et al., 2008), identifies synapses using GFP complementation of the NRX-NLG interaction. A fragment of GFP is fused to the ectodomain of NRX while the complementary GFP fragment is fused to the ectodomain of NLG. Formation of a NRX-NLG adhesion complex at synapses recombines the GFP fragments, and fluorescence is restored an hour or more later. The primary limitations of GRASP for NRX-NLG interaction detection are that GFP recombination is irreversible (Feinberg et al., 2008; Yamagata and Sanes, 2012) and GFP fluorescence is dim. The irreversibility can shift the equilibrium between the complexed and non-complexed states of NRX-NLG, and preclude dynamic reporting of NLG-NRX interactions upon stimulation.

In 2010, the Ting lab presented an alternative approach to image trans-synaptic NRX-NLG interactions based on enzymatic biotinylation of an acceptor peptide (AP) by *E. coli* biotin ligase (BirA, Figure 6-1) (Thyagarajan and Ting, 2010) (paper now retracted). In that report, AP was fused to NLG and BirA was fused to NRX. When expressed in different but contacting neurons, site-specific biotinylation, detected by staining of live neurons with streptavidin-fluorophore conjugates, was reported at synaptic contacts. This method was named BLINC, for Biotin Labeling of INtercellular Contacts (Thyagarajan and Ting, 2010). Since this publication, we discovered that the work in this paper cannot be reproduced, and wanted to examine the technical reasons for irreproducibility, make changes in the BLINC constructs and protocols in order to achieve successful BLINC labeling in neuronal cultures, and then introduce an improved method for NRX-NLG contact imaging based on lipoic acid ligase instead of biotin ligase (Fig. 6-1).
Figure 6-1. Scheme showing BLINC and ID-PRIME methods for imaging trans-cellular protein-protein interactions.
In Biotin Labeling of Intercellular Contacts (BLINC), protein A is genetically tagged with the 35 kDa *E. coli* biotin ligase (BirA) on the extracellular side. Protein B is genetically tagged with a 15-amino acid acceptor peptide (AP) for BirA. When proteins A and B interact, BirA ligates biotin onto protein B, which can be detected using a monovalent streptavidin-fluorophore conjugate (Howarth et al., 2006a).

We started by applying the constructs from the 2010 work (Thyagarajan and Ting, 2010) in human embryonic kidney 293T (HEK) cells (Figure 6-2). Two pools of HEK cells were separately transfected with the BirA$_{64}$-NRX1β fusion (numbering indicates the BirA insertion site – at amino acid 64 of the immature NRX1β protein in this case) and the AP-NLG1 fusion. The two HEK populations were then resuspended, plated together, and allowed to form contacts over 24 hours. Trans-cellular biotinylation was initiated with the addition of biotin-AMP ester (Thyagarajan and Ting, 2010) for 2 min. Sites of AP biotinylation were detected on living cells by staining with streptavidin-AF568 conjugate (Figure 6-2A). Images in Figure 6-2B show biotinylation sites (BLINC signal) localized to NRX-NLG contacts, as indicated by the YFP and BFP co-transfection markers. AP-NLG1-expressing cells not contacting BirA cells were not labeled.

Figure 6-2. BLINC for imaging neurexin-neuroligin interactions in HEK cells.
(A) Scheme showing the BLINC experimental protocol. Two pools of HEK cells were separately transfected with BirA$_{64}$-NRX1β plus YFP, or AP-NLG1 plus BFP. The pools were then mixed and allowed to form contacts over 24 hours. BLINC labeling was performed with 10 μM biotin-AMP for 2 minutes (note that biotin+ATP was used instead for neuron cultures in other figures, for reasons...
explained in Fig. 6-9). Biotinylated AP sites were detected by live-cell staining with streptavidin-AF568 for 5 minutes. (B) BLINC imaging results. Controls are shown with a D137A mutation in BirA-NRX to abolish its interaction with NLG (rows 2 and 4), and 1 μM exogenous BirA added during the biotin-AMP step to label total cell surface AP-NLG1 (rows 3 and 4). When a NLG-expressing cell apposes a NRX-expressing cell, BLINC signal is localized at contact sites (thin arrow heads, row 1). The same phenomenon was observed when exogenous BirA was added to label the total NLG pool (thick arrow heads, row 3). All scale bars, 10 μm.

To test if BLINC labeling was interaction-dependent, we introduced a point mutation (D137A) in NRX to abolish Ca$^{2+}$ binding (Graf et al., 2004) and therefore eliminate trans-interaction with NLG1. Figure 6-2B shows that the mutant construct, BirA$_{64}$-NRX1β (D137A), gave almost no detectable BLINC staining at contact sites with AP-NLG1-expressing cells. As a positive control, we used exogenous BirA (purified BirA enzyme added to the cell media) to biotinylate the total surface pool of AP-NLG1, regardless of its proximity to a BirA-NRX1β-expressing cell. The third row in Figure 6-2B shows streptavidin staining of all AP-NLG1 expressing cells, not only those in contact with BirA-NRX1β-expressing cells. Interestingly, for blue cells in contact with green cells, the streptavidin signal was still localized to cell-cell contact sites, suggesting that when NRX and NLG expression levels are matched, their binding affinity is strong enough to aggregate the total surface protein pools at these contact regions. A similar control with BirA-NRX1β (D137A)-expressing cells also showed labeling of all AP-NLG1-expressing cells by exogenous BirA (fourth row), but the streptavidin signal was not localized to contact sites between green and blue cells. We believe that AP-NLG1 distributes evenly around the perimeter of the transfected cell because the NRX1β (D137A) mutant is unable to trap AP-NLG1 at contact sites. From this experiment, we conclude that NRX-NLG BLINC is robust and reproducible in HEK cells. Unfortunately, we found that this was not the case in neuron cultures. After many unsuccessful efforts to reproduce neuron BLINC using the previously described NRX and NLG fusion constructs and protocols (Thyagarajan and Ting, 2010), we decided to systematically examine the fundamental aspects of the system.

Problems with the CMV-AP-NLG1 Construct from Reference (Thyagarajan and Ting, 2010)
The design of the BLINC reporter system is such that AP and BirA fusions must be introduced into separate but contacting cells. HEK cells can be separately transfected, then lifted and replated together, but neurons cannot be replated without damaging their delicate processes and synapses. Therefore, it is necessary to transf ect them immediately after dissociation, while they are still in suspension, and then plate them together thereafter. We previously opted for nucleofection-type transfection (Thyagarajan and Ting, 2010), because it is compatible with suspended neurons, gives high transfection efficiencies (necessary in order to see a reasonable number of overlapping BirA- and AP-containing processes), and eliminates the possibility of plasmid overlap, where BirA and AP fusions express together in the same neuron.

We first examined the expression of BLINC constructs in neurons by introducing the AP-NLG1 construct alone, using nucleofection, into suspended hippocampal neurons at 0 days in vitro (DIV0). Neurons were then plated, and five days later (since previous experiments were all reported at DIV5 and DIV16), we checked for expression by performing exogenous biotinylation with purified BirA added to the culture medium. This assay is expected to give a much stronger signal than any BLINC experiment, because BirA is provided in great excess, and total AP rather than just synaptic AP will be biotinylated. Figure 6-3 shows that no biotinylation was detected. We also found that biotinylation was undetectable at DIV12. We were unable to check at DIV16, because in order to see overlapping transfected processes at DIV5, it was necessary to plate neurons at a high initial density. As a result, neurons were often dense and unhealthy at DIV16, making it difficult to distinguish specific biotinylation from non-specific binding of streptavidin conjugates to unhealthy cells. This published AP-NLG1 construct (Thyagarajan and Ting, 2010) used a CMV promoter, which, according to previous reports, may give inconsistent (Kaech and Banker, 2006) and activity-dependent expression in transiently transfected neuron cultures (Wheeler and Cooper, 2001). We therefore created an identical construct driven by the CAG promoter (Niwa et al., 1991) (CMV enhancer/chicken β-actin) instead. This promoter has been used previously for strong transgene expression.
in neurons (Chih et al., 2006), (Siddiqui et al., 2010). Using CAG-AP-NLG1, we were able to detect weak but specific biotinylation in some neurons at DIV5 and DIV12 (Figure 6-3).

In another effort to improve the signal, we used a construct with three tandem AP tags: CAG-3xAP-NLG1. Figure 6-3 shows that exogenous biotinylation of neurons nucleofected with this construct produced signal well above background and specific to transfected cells, at both DIV5 and DIV12. All these comparisons were performed in parallel under identical conditions. We conclude that the CAG rather than CMV promoter was essential to give long-lasting (>5 day) expression of this fusion construct in transiently transfected neuron cultures, and the 3xAP rather than 1xAP tag on NLG was necessary to give streptavidin signal above background in the majority of neurons at both DIV5 and DIV12. The previous CMV-AP-NLG1 (Thyagarajan and Ting, 2010) construct lacked both persistent expression and detectability above background, which explains why it cannot be successfully used for BLINC in neurons.

**Figure 6-3.** Expression of CMV-AP-NLG1 cannot be detected in neurons, but CAG promoter constructs can be detected. The indicated plasmids were introduced by nucleofection, along with a Venus marker (shown in green), into DIV0 dissociated rat hippocampal neurons. At either DIV5 (left), or DIV12 (right), surface AP fusion proteins were labeled with 1 uM exogenous BirA (+ biotin and ATP), followed by streptavidin-AF647 (shown in red), then imaged live. Streptavidin channel intensities are normalized within the DIV5 dataset and within the DIV12 dataset, but not across datasets. Labeling of CMV-1xAP-NLG1 was not detected across 10 transfected cells at DIV5, and 23 transfected cells at DIV12. Labeling of CAG-1xAP-NLG1 was detectable but weak in 4 of 10 transfected cells at DIV5, and 12 of 23
transfected cells at DIV12. Labeling of CAG-3xAP-NLG1 (with three AP tags in tandem) was generally stronger, and detected in 8 of 9 neurons at DIV5, and 12 of 16 transfected neurons at DIV12. On the right, the arrowhead points to a lightly streptavidin-labeled cell that expressed the Venus marker weakly. Scale bars, 10 μm.

Problems with the CMV-BirA-NRX1β Construct from Reference (Thyagarajan and Ting, 2010)

We next turned our attention to the BirA-NRX fusion construct. We introduced the previously published construct, CMV-BirA$_{64}$-NRX1β (Thyagarajan and Ting, 2010), by nucleofection into DIV0 hippocampal neurons. Figure 6-4 shows that expression was not detected by anti-c-Myc staining at both DIV5 and DIV12, although positive controls with the same construct introduced one day before labeling, by lipofection instead, were detectable.

![Figure 6-4. Expression of CMV-BirA$_{64}$-NRX1β from reference (Thyagarajan and Ting, 2010) cannot be detected in neurons after nucleofection.](image)

The CMV-BirA$_{64}$-NRX1β plasmid was introduced by nucleofection, along with a Venus marker (shown in green), into DIV0 dissociated rat hippocampal neurons. Anti-c-Myc staining was performed on living cells to detect surface expression of the BirA-NRX at DIV5 (top) and DIV12 (bottom). As a positive control, staining was performed in parallel on neurons transfected with the same plasmid, using lipofectamine instead of nucleofection, 1 day before the labeling experiment. In general, we find that lipofection of a plasmid gives much higher expression in neurons than nucleofection of the same plasmid. AF647 channel intensities are normalized within DIV5 and DIV12 datasets, but not across datasets. For samples nucleofected with CMV-BirA$_{64}$-NRX1β following the protocol in reference (Thyagarajan and Ting, 2010).
and Ting, 2010), c-Myc staining could not be detected across 13 transfected cells at DIV5 and 16 transfected cells at DIV12. For lipofected control samples, c-Myc staining could be detected in 4 of 5 cells at DIV5, and 4 of 7 cells at DIV12. In general, c-Myc staining on lipofected neurons was weaker at DIV12 than at DIV5 in this experiment. Scale bars, 10 μm.

We reasoned again that the CMV promoter could be part of the problem, so we switched to a CAG promoter. Figure 6-5 shows that CAG-BirA-NRX could be detected 4 days after lipofection of neuron cultures whereas CMV-BirA-NRX could not. Based on these experiments, we conclude that the CAG promoter is essential to give persistent expression of this construct as well, and the previously published CMV-BirA64-NRX1β (Thyagarajan and Ting, 2010) was undetectable in neuron cultures five days after its introduction by nucleofection. This explains why it cannot be successfully used for BLINC in neurons.

![Figure 6-5. The CAG promoter gives more persistent expression of BirA64-NRX1β in neurons than the CMV promoter. BirA64-NRX1β with either a CMV or CAG promoter was introduced to plated hippocampal neurons at DIV5 using lipofectamine, along with a synaptophysin-YFP marker (shown in green). Expression was detected 4 days later, at DIV9, by anti-c-Myc staining on either living neurons (top) or fixed and permeabilized neurons (bottom). c-Myc staining background was very high for fixed neurons. Whereas CMV-BirA64-NRX1β expression could not be detected across multiple fields of view, CAG-BirA64-NRX1β expression was detectable 4 days after lipofection. Note that in Figure 6-4, lipofected CMV-BirA64-NRX1β was detected 1 day rather than 4 days after lipofection. Scale bars, 10 μm.](image-url)
Identifying optimal sites for fusing biotin-ligase in neurexin

Using the improved CAG-BirA-NRX1β and CAG-3xAP-NLG1 constructs that give persistent expression in neurons after nucleofection, we attempted BLINC in neurons again, but were still unsuccessful. Based on the images in Figure 6-5, we suspected that part of the problem might be the poor trafficking of the BirA-NRX fusion to the cell surface and to synapses; the majority of it appeared to be intracellular and localized to the cell body rather than distal processes. Indeed, immunofluorescence staining in HEK cells (Figure 6-6C) showed that BirA64-NRX1β was mostly trapped in the secretory pathway compared to c-Myc-LAP-NRX1β (tag size 35 kD vs. 2.6 kD), suggesting that the large BirA tag disrupted trafficking, and that its insertion site would have to be optimized. Previous studies have inserted large tags, such as fluorescent proteins, into the cytosolic tail of NRX (Graf et al., 2004), (Fairless et al., 2008) or into its stalk domain (Taniguchi et al., 2007), (Fu and Huang, 2010), an extracellular region proximal to the transmembrane segment.
Figure 6-6. Trafficking of BirA fusion constructs in HEK and neurons. (A) Domain structures of BirA and AP fusions to NRX3β, NRX1β, and NLG1 used in this figure. TM is the transmembrane domain. (B) BirA and AP insertion sites in NRX and NLG. A side-on view into the synaptic cleft is shown for the dimeric extracellular domain of NLG1 (amino acids 52-634) in
complex with two extracellular domains of NRX1β (amino acids 82-288; colored orange). From PDB 3VKF (Tanaka et al., 2012). Ca²⁺ ions are shown in green. Note that amino acid 288 of NRX1β corresponds to amino acid 259 of NRX3β. (C) Trafficking in HEK cells. Cells were transfected with the indicated constructs, fixed and permeabilized, then stained with the indicated antibodies. Fluorescence images are not normalized. Bottom row shows overlay onto DIC images. The BirA tag reduces surface trafficking of NRX1β, NRX3β, and NLG1. (D) Trafficking in neurons. Hippocampal neurons were lipofected at DIV11 with the indicated constructs and a Venus co-transfection marker (shown in green). One day later, neurons were stained live with anti-HA antibody to visualize surface expression. (E) Same as (D) except that neurons were fixed and permeabilized before staining with anti-HA antibody to visualize total protein pools, rather than surface pools only. Trafficking of NRX3β and NLG1 to processes is impaired when either is fused to BirA. Scale bars, 10 μm.

We wondered if moving the BirA tag to different locations might improve the surface targeting of our NRX fusion. We prepared two new extracellular fusions of BirA to the NRX3β gene. NRX3β is in many ways functionally interchangeable with NRX1β; the two isoforms display similar endogenous localization in neurons, possess similar trans-cellular binding affinity to NLG1 (Koehnke et al., 2010a), and their crystal structures can be overlaid without significant differences (Araç et al., 2007; Koehnke et al., 2010b). Figure 6-6C shows immunofluorescence staining of BirA36-NRX3β (N-terminal fusion, after signal peptide) and BirA272-NRX3β (stalk domain fusion), compared to HA-NRX3β, in HEK. Again, the BirA fusions were impaired in their surface targeting compared to HA-NRX3β, where the latter was predominantly localized to the cell surface. We also prepared a BirA fusion to NLG1 at its N-terminus after the signal peptide (BirA48-NLG1) and found that it too was largely trapped inside the cell compared to HA-AP-NLG1 (tag size 35 kD vs 2.9 kD) (Figure 6-6C).

Nevertheless, we tested our three new BirA fusion constructs in neurons. Figure 6-6D shows live-cell immunostaining of DIV12 neurons transfected with each construct. HA-NRX3β produced a very strong signal specific to transfected neurons. BirA36-NRX3β and BirA272-NRX3β were much weaker, but still detectable above background. In contrast, surface expression of BirA48-NLG1 was undetectable. Figure 6-6E shows the same experiment but with immunofluorescence staining performed after neuron fixation to detect total protein pools. From these images it was apparent that HA-NRX3β and AP-NLG1 proteins could be found in distal processes, while the BirA fusions were
predominantly localized to the cell bodies. We concluded that while none of our fusion sites tolerate the 35 kD BirA tag well, the NRX fusions are better than the NLG fusion, so we proceeded to try BLINC experiments with these.

**Biotin-ligase dependent labeling of the neurexin-neuroligin complex in neurons**

First, we verified that both BirA\textsubscript{36}-NRX3\(\beta\) and BirA\textsubscript{272}-NRX3\(\beta\) gave detectable and localized trans-cellular BLINC labeling with 3xAP-NLG1 in HEK cultures (data not shown), demonstrating that they are functionally competent. Then we performed a BLINC experiment in hippocampal neurons, introducing each construct into separate pools of suspended DIV0 neurons by nucleofection. Figure 6-7A shows that whereas BirA\textsubscript{272}-NRX3\(\beta\) produced BLINC labeling at sites of overlap with 3xAP-NLG1-expressing neurons (indicated by overlap of green and blue transfection markers), BirA\textsubscript{36}-NRX3\(\beta\) did not. This trend was observed across >15 fields of view in this experiment. The observation that the site of BirA insertion into the extracellular domain of NRX3\(\beta\) influenced BLINC sensitivity is interesting in light of the fact that no difference in signal between these two constructs was seen in HEK cultures (data not shown). Perhaps the presence of endogenous NRX interaction partners in neurons alters NRX’s conformation and decreases the steric accessibility of fused BirA\textsubscript{36}, but not BirA\textsubscript{272}. The immunofluorescence controls in Figure 6-6D show that the difference in BLINC outcomes in neurons cannot be explained by a difference in surface expression levels for BirA\textsubscript{36}-NRX3\(\beta\) versus BirA\textsubscript{272}-NRX3\(\beta\).

We noticed that the BLINC signal in neurons was strongest at sites of overlap between transfected cells, but streptavidin staining was also detected at non-overlapping regions on the 3xAP-NLG1-expressing cell (Figure 6-7A), suggesting that AP-NLG biotinylated in *trans* by BirA could subsequently diffuse away from the site of interaction. This contrasted with our observations in HEK cells, where the BLINC signal was tightly localized to cell-cell contacts and not
diffusive (Figure 6-2B). We hypothesized that this discrepancy in localization of BLINC signal resulted from a difference in BirA:AP stoichiometry in these two experimental configurations. Since the insertion of BirA into NRX3β strongly impeded NRX3β trafficking to the surface of neurons while AP-NLG1 did not have an apparent trafficking defect (Figure 6-6D–E), there was probably insufficient BirA-NRX3β to anchor biotinylated 3xAP-NLG1, and the latter could diffuse away from the contact site after biotinylation. This explanation is supported by our observations in a mixed culture experiment, where HEK cells expressing BirA272-NRX3β were plated on top of hippocampal neurons expressing 3xAP-NLG1 (Figure 6-7B). Here, the resulting BLINC signal tightly localized to contact sites and did not diffuse outward. We believe this is because the quantity of BirA-NRX3β presented on the surface of the overlaid HEK cell was much higher than that presented on the surface of an overlaid neuron, and therefore anchoring of the biotinylated 3xAP-NLG1 pool could occur. Using this mixed culture assay, we also performed a negative control with the non-interacting D137A mutant of NRX and observed an absence of BLINC signal (Figure 6-7B).
Figure 6-7. BLINC for imaging neurexin-neuroligin interactions in neuron cultures and in HEK-neuron mixed cultures.

(A) BLINC labeling of pure neuron cultures. Two pools of hippocampal neurons were separately nucleofected at DIV0 with BirA-NRX plus a membrane tdTomato marker (shown in blue), or 3xAP-NLG1 plus a Venus marker (shown in green). For the top row, the BirA36-NRX3p construct was used, and for the bottom row the BirA272-NRX3p construct was used. All constructs had CAG promoters. Labeling was performed at DIV5 with biotin+ATP for 15 minutes, followed by monovalent streptavidin-AF647 detection for 5 minutes. Confocal images of live neurons showed no detectable BLINC signal for the BirA36-NRX3P fusion across 10 fields of view in which Venus- and Tomato-expressing neurons were observed to be crossing. For the BirA272-NRX3p fusion (bottom row), BLINC signal was detected in 5 out of 10 such fields of view.

(B) BLINC labeling of mixed HEK-neuron cultures. HEK cells expressing BirA272-NRX3P and a dsRed marker (shown in blue) were plated on top of rat hippocampal neurons transfected with lipofectamine at DIV10 with 3xAP-NLG1 plus a Venus marker (shown in green). Labeling was performed at DIV11 as in (A). BLINC signal could be detected in 22 out of 30 fields of view, and was localized to contact sites (arrow heads). The bottom row shows a control with a D137A mutation in NRX3β; BLINC signal was not observed in any field of view. All scale bars, 10 μm.

Optimization of BLINC Labeling Reagent

Another feature of our BLINC signal in neurons (Figure 6-7A) is that it is clearly not synaptic. Apposing neurons differentially expressing the transgenic NRX and NLG fusions sometimes “zipped up” along one another’s processes, establishing large zones of contact that were clearly not synapses. This is likely
an overexpression artifact. Interestingly, when neurons were transfected with either BLINC construct alone, the constructs displayed relatively decent overlap with pre- and post-synaptic markers (Figure 6-8), but in a trans experiment, the affinity of the overexpressed fusion constructs for one another greatly perturbed neuron morphology.

**Figure 6-8. Synaptic localization of optimized BLINC constructs in neurons.** Hippocampal neurons were lipofected with the indicated constructs and synaptic markers at DIV11 and imaged live at DIV12. The pre-synaptic marker synaptophysin-YFP was used at left, and the post-synaptic marker Homer-GFP was used at right, both shown in green. Anti-HA staining was performed on living neurons to visualize surface pools of BirA (left) and AP (right) fusion proteins. In the “merge” panel, yellow indicates sites of red-green overlap. All scale bars, 10 μm.

Since BirA has a higher affinity for the biotin-AMP intermediate than ATP (Kwon and Beckett, 2000), biotin-AMP can be supplied at micromolar concentrations and still produce labeling signal comparable to millimolar concentrations of ATP (+20 μM biotin) (Howarth and Ting, 2008). Initially, as per the reported protocol (Thyagarajan and Ting, 2010), we attempted BLINC labeling in neurons using biotin-AMP in place of biotin+ATP (Figure 6-9). This option is less likely to activate endogenous purine receptors on neurons and cause toxicity (Rathbone et al., 1999). We observed, however, that even 2.5 μM biotin-AMP resulted in much higher background fluorescence on neurons than 20 μM biotin +1 mM ATP (Figure 6-9). Higher biotin-AMP concentrations worsened the background without increasing the BLINC signal. This problem was not observed in BLINC experiments in HEK cells (Figure 6-2B), so we surmise that the high-energy biotin-AMP reagent was covalently reacting with nucleophiles on polylysine/laminin-coated coverslips. Our observations are inconsistent with the previous BLINC study, in which 10 μM biotin-AMP was used without detectable background. We note, however, that biotin-AMP can be used successfully for biotinylation of total surface AP-NLG1 pools on neurons, as in reference
(Howarth et al., 2006b), because the signal is so much stronger than BLINC signal that it can be clearly detected above the biotin-AMP-related background.

**Figure 6-9. Use of biotin-AMP for BLINC in neuron cultures generates high imaging background.** Hippocampal neurons were nucleofected at DIV0 with BirA\textsubscript{272}−NRX3β plus dsRed (shown in blue), or 3xAP-NLG1 plus Venus (shown in green). The two pools were mixed together and allowed to form contacts. At DIV9, cells were labeled with biotin+ATP, or biotin-AMP ester, as indicated for 5 minutes, then stained with monovalent streptavidin-AF647 (shown in red) for another 5 minutes and imaged live. On the right are images of untreated coverslips. From this experiment we conclude that signal intensities are similar for biotin+ATP, and 2.5 \textmu M biotin-AMP. However, the nonspecific background is higher when using 2.5 \textmu M biotin-AMP. In contrast, the background when using biotin+ATP is the same as for untreated coverslips, i.e., undetectable. Note that the problem of high background with biotin-AMP is observed only for neurons, and is not seen when performing BLINC or exogenous BirA labeling on HEK cells (as in Figure 6-2B). Scale bars, 10 \textmu m.

**NRX-NLG BLINC in Neurons is not Robust**

Even with our improvements to the BLINC reporter constructs and labeling protocol, we found that BLINC labeling in neurons was not robust. BLINC signal intensity varied from nucleofection to nucleofection and sometimes was not detectable at all. We believe this is a consequence of variations in BirA-NRX surface expression levels – sometimes it fell below the threshold necessary to produce detectable labeling signal. We conclude that BLINC in its current form can be a powerful and robust tool for imaging NRX-NLG interactions in HEK cultures (as in Figure 6-2B) and in HEK-neuron mixed cultures (as in Figure 6-
but the technology in its current form is too unreliable in pure neuron cultures. For this reason, we turned our attention to an alternative methodology for NRX-NLG contact imaging in neurons.

**Lipoic acid ligase (LpIA) dependent labeling of the neurexin-neuroligin complex**

![Scheme showing ID-PRIME for imaging trans-cellular protein-protein interactions.](image)

**Figure 6-10. Scheme showing ID-PRIME for imaging trans-cellular protein-protein interactions.**

In Interaction-Dependent PRobe Incorporation Mediated by Enzymes (ID-PRIME), protein A is genetically tagged with a 38 kDa mutant of *E. coli* lipoic acid ligase (*LpIA=W37A, T57I, F147L, H267R* mut of LpIA) on its extracellular side. Protein B is genetically tagged with a 13-amino acid ligase acceptor peptide (LAP) for LpIA. When proteins A and B interact, *LpIA* ligates lipoic acid onto protein B, which can be detected using an antibody-fluorophore conjugate as shown in the top row. Bottom row show alternative ID-PRIME detection using picolyl azide ligation onto protein B. Ligated azide can be detected by copper-catalyzed click chemistry with alkyne-fluorophore conjugates (Uttamapinant et al., 2012).

The Ting lab has previously developed a suite of methods for targeting chemical probes to specific proteins in living cells using engineered mutants of *E. coli* lipoic acid ligase (LpIA). These methods are collectively called "PRIME", for PRobe Incorporation Mediated by Enzymes (Uttamapinant et al., 2012), (Baruah et al., 2008a; Liu et al., 2012a; Slavoff et al., 2011; Uttamapinant et al., 2010; Yao et al., 2012). PRIME works in a similar way to BirA-mediated biotinylation, with the ligase catalyzing covalent conjugation of a small molecule to a recognition peptide (the Ligase Acceptor Peptide, or LAP), but our engineering of the LpIA active site has made it possible to conjugate a wide range of chemical...
structures besides lipoic acid, including fluorophores (Uttamapinant et al., 2010), photocrosslinkers (Baruah et al., 2008b), and functional group handles (Yao et al., 2012). We have also found that the LpIA/LAP pair can be used for detection of cytosolic protein-protein interactions by ID-PRIME (Interaction-Dependent PRIME), when the affinity of LAP for LpIA is tuned such that probe ligation occurs only when the proteins to which LpIA and LAP are fused interact (Slavoff et al., 2011). We wondered if the LpIA/LAP pair could be used for detection of intercellular protein-protein interactions in a manner analogous to BLINC, as shown in Figures 6-10.

Figure 6-11. Comparison of surface trafficking in neurons for BLINC and ID-PRIME ligase fusion constructs. (A) Domain structures of LpIA, BirA, and LAP fusion constructs used in this figure and Figures 6-12 & 6-13. TM is the transmembrane domain. HA tags are colored red and a linker is colored green. (B) Comparison of surface trafficking in neurons for BLINC and ID-PRIME ligase fusion constructs. Hippocampal neurons were nucleofected at DIV0 with *LpIA$_{36}$-NRX3β or BirA$_{272}$-NRX3β, plus a membrane tdTomato marker (shown in green). At DIV5, surface expression of each construct was detected by live-cell immunostaining with anti-HA antibody, shown in red at two different intensity levels. *LpIA$_{36}$-NRX3β surface expression was easily detected in 19 out of 19 transfected neurons, while BirA$_{272}$-NRX3β surface expression was undetectable in 10 out of 10 transfected neurons. Note that in Figure 6-6D, surface detection of BirA$_{272}$-NRX3β was performed after lipofection, not nucleofection. Scale bars, 10 μm.

There were a few considerations before we could attempt such an experiment. First, we previously observed that LpIA and its mutants have high activity in the mammalian cytosol, but the activity drops for unknown reasons when LpIA is targeted to the secretory pathway or the cell surface (Uttamapinant
et al., 2010). Separate efforts in our lab have produced, using yeast display evolution, a quadruple mutant of LpIA with higher activity in the secretory pathway and on the cell surface (White and Zegelbone, 2013a). Second, we considered which LAP sequence to use: the regular, high affinity sequence used for most PRIME experiments with a $K_M$ of 13 $\mu$M (Puthenveetil et al., 2009), or the lower affinity sequence used for intracellular ID-PRIME with a $K_M >$200 $\mu$M(Slavoff et al., 2011). We opted for the high affinity sequence because we predicted that the lower effective protein concentrations in a trans-cellular experiment would render PRIME still interaction-dependent – just as we observed BLINC to be interaction-dependent with regular AP ($K_M$ of 25 $\mu$M (Beckett et al., 2008)) (Figure 6-2B) even though a modified, lower-affinity AP(-3) peptide was previously necessary for interaction-dependent biotinylation by BirA in the cytosol (Fernández-Suárez et al., 2008). Third, we considered which of many PRIME probes to use for trans-interaction readout at the cell surface. We selected the natural substrate lipoic acid (Figure 6-10) and picolyl azide (Figure 6-10). The former, detectable by antibody-fluorophore conjugates, is advantageous for its superior ligation kinetics compared to unnatural substrates (Liu et al., 2012b; Uttamapinant et al., 2010; Yao et al., 2012). The latter is attractive because detection of the picolyl azide is performed entirely with small-molecule reagents ("click" chemistry with alkyne-fluorophore conjugates (Uttamapinant et al., 2012)), which have better steric access to crowded cellular junctions, do not induce crosslinking, and minimize perturbation to the subsequent trafficking and internalization of labeled proteins compared to detection by antibodies or streptavidin.

**ID-PRIME in HEK Cells**

Three tandem LAP tags (3xLAP) were introduced onto the N-terminus of NLG1. A 38 kD LpIA mutant evolved by Dr. Katherine White with improved activity in the secretory pathway (mutations: W37A, T57I, F147L, H267R (White and Zegelbone, 2013b), referred to below as *LpIA) was fused to the N-terminus of NRX3β after amino acid 36. Lipoic acid ID-PRIME was successfully performed
in HEK cultures, with antibody signal detected between transfected cells (Figure 6-12A). Negative controls with lipoic acid omitted or the LAP tag replaced by AP produced no signal. Like BLINC, ID-PRIME labeling was interaction-dependent because a NRX3β D137A mutation in the *LpIA₃β-NRX3β construct eliminated labeling (Figure 6-12A, bottom row).

Figure 6-12. ID-PRIME for imaging neurexin-neuroligin interactions in HEK cells. (A) ID-PRIME with lipoic acid readout (as in Figure 6-10). HEK cells were separately transfected with *LpIA₃β-NRX3β plus a membrane-localized tdTomato marker (shown in blue), or 3xLAP-NLG1 plus a Venus marker. After mixing and replating, cells were labeled with 50 µM lipoic acid.
+500 μM ATP for 15 minutes. Ligated lipoic acid was detected with an anti-lipoic acid antibody followed by a secondary antibody-AF647 conjugate (shown in red) for 5 minutes each. For row 1, a magnified view representing the boxed region, and a more contrasted view of the transfection markers are shown on the right. Controls were performed with lipoic acid omitted (row 2), the acceptor peptide for BirA substituted for LAP (row 3), and the interaction-deficient NRX mutant (row 4). (B) ID-PRIME with picolyl azide readout (as in Figure 6-10). HEK cells were transfected as in (A), and labeling was performed with 100 μM picolyl azide +500 μM ATP for 15 minutes, followed by detection with copper-catalyzed click chemistry, using 50 μM copper and 20 μM alkyne-AF647. Color schemes and controls are the same as for (A). All scale bars, 10 μm.

Picolyl azide ID-PRIME using these same fusion constructs was also successfully performed in HEK cultures (Figure 6-12B). Again, negative controls with azide omitted, LAP replaced by AP, or a D137A mutation in NRX3β showed no signal. Here, the ID-PRIME signal (from Alexa Fluor 647-alkyne) was more clearly concentrated at junctions between LpIA- and LAP-expressing cells. This is probably because the small molecule detection reagents for picolyl azide ID-PRIME could better access the crowded adhesion junctions compared to antibody detection reagents used for lipoic acid ID-PRIME.

We also compared 1xLAP-NLG1 to 3xLAP-NLG1 for lipoic acid ID-PRIME and found that the tandem LAPs did not boost signal as strikingly as tandem APs did (data not shown), possibly because the 3xLAP tag reduced NLG1 expression at the surface (compared to 1xLAP), or because tandem lipoic acid molecules in close proximity could not be simultaneously accessed by antibodies.
Figure 6-13. ID-PRIME for imaging neurexin-neuroligin interactions in neuron cultures and in HEK-neuron mixed cultures. (A) Lipoic acid ID-PRIME labeling of pure neuron cultures. Dissociated rat hippocampal neurons were separately nucleofected at DIV0 with either 1xLAP-NLG1 plus a Venus transfection marker (shown in green), or *LpIA36-NRX3β plus a membrane-localized tdTomato transfection marker (shown in blue). The two pools of neurons were mixed and plated. At DIV5, neurons were labeled with lipoic acid and anti-lipoic acid antibody as in Figure 6-11A. ID-PRIME signal was detected in 22 out of 23 fields of view, and was localized to contact sites (arrow heads, row 1). Negative controls with lipoic acid omitted (row 2), AP-NLG1 in place of LAP-NLG (row 3), or with an interaction deficient mutant of NRX (row 4) are also shown. Asterisks in row 1 and 2 indicate sites where the over-expression of ID-PRIME constructs caused neuronal processes to "zip up". (B) Lipoic acid ID-PRIME labeling of mixed HEK-neuron cultures. HEK cells expressing *LpIA36-NRX3β and a membrane-localized tdTomato marker (shown in blue) were plated on top of neurons, transfected with lipofectamine at DIV7 with 3xLAP-NLG1 plus a Venus marker (shown in green). Labeling was performed as in Figure 6-11A, at DIV8. ID-PRIME signal was detected in 9 out of 11 fields of view in which Venus-expressing neurons contacted Tomato-expressing HEK cells. The bottom row shows a control with a D137A mutation in NRX; no ID-PRIME signal was observed in any field of view. All scale bars, 10 μm.

Lipoic Acid ID-PRIME in Neurons

We proceeded to test labeling in neuron cultures. Using the same nucleofection protocol developed for BLINC, lipoic acid ID-PRIME signal was detected at overlap sites between neurons expressing *LpIA36-NRX3β and neurons expressing 1xLAP-NLG1 (Figure 6-13A). Controls with lipoic acid omitted, LAP replaced by AP, or a D137A mutation in NRX eliminated the signal. In contrast to BLINC, ID-PRIME signal was localized to overlapping sites and did not appear to spread outward on the LAP-expressing neuron. We believe this is because the LpIA-NRX:LAP-NLG stoichiometry was better matched than the BirA-NRX:AP-NLG stoichiometry in these neuron experiments. This is supported by the observation that *LpIA36-NRX3β surface expression in neurons after nucleofection was much higher than surface expression of our best BLINC construct, BirA272-NRX3β, under identical conditions (Figure 6-11). This higher expression also helps to explain why lipoic acid ID-PRIME labeling was much more robust and reproducible than BLINC labeling in neurons.
We quantified the sensitivity of lipoic acid ID-PRIME in neurons and found that 38–54% of contact sites between transfected neurons (defined by the intersection of fluorescent protein transfection markers for *LpIA<sub>36</sub>-NRX3β and 1xLAP-NLG) contained ID-PRIME signal. The lack of signal at ~46% of contacts could be due to sensitivity limits of the methodology, or an absence of true NRX-NLG interactions at fluorescent marker intersection sites. Like in our neuron BLINC experiments, ID-PRIME constructs also induced “zipping” of neuronal processes (Figure 6-13A, marked sites in rows 1 and 2) when overexpressed.

Lipoic acid ID-PRIME also detected the trans-cellular interaction between 3xLAP-NLG1 expressed in neurons and *LpIA<sub>36</sub>-NRX3β expressed in overlaid HEK cells (Figure 6-13B). A D137A mutation in NRX eliminated ID-PRIME signal in this mixed culture configuration.

**ID-PRIME by double lipofection**

To introduce reporter plasmids into separate pools of neurons by nucleofection is labor intensive and consumes large numbers of neurons. Lipofection of plated neurons (Dalby et al., 2004) is much simpler and uses fewer cells, so we also attempted lipoic acid ID-PRIME using this strategy. Since lipofection usually transfects a somewhat random and small subset of neurons, and the lipofecting medium can be removed within hours of addition, we surmised that it would be possible, by sequential lipofection, to generate a culture in which some neurons express NRX only, some neurons express NLG only, and some express both. Figure 6-14 shows neuron cultures lipofected with *LpIA<sub>36</sub>-NRX3β at DIV5, and again with LAP-NLG1 one day later, at DIV6. After labeling and antibody staining at DIV7, we detected trans signal in many fields of view (evident from ID-PRIME signal on top of overlapping green and blue processes), but contaminating cis signal from neurons co-expressing both fusion constructs was also observed in some fields of view. The cis signal was generally much stronger than the trans signal, and spread over the entire surface of the
transfected neuron, instead of being localized to green-blue overlap sites. We conclude that this experimental configuration is useful and much easier to implement, but one must be cautious about interpreting signal as being trans or cis in origin.

![Diagram showing signal localization](image)

**Figure 6-14.** Lipoic acid ID-PRIME with lipofected neuron cultures. Same as Figure 6-12A, except that constructs were introduced by sequential lipofection into plated hippocampal neurons at DIV5 and DIV6, instead of by nucleofection into separate pools of DIV0 neurons (which ensures complete plasmid segregation). For lipofection, *LpIA-NRX* plus a membrane tdTomato marker (shown in blue) were first introduced at DIV5, then the same cultures were lipofected again at DIV6 with 1xLAP-NLG1 plus a Venus marker (shown in green). All constructs had CAG promoters. At DIV7, lipoic acid ID-PRIME labeling was performed as in Figure 6-11A. Expression of *LpIA-NRX* and 1xLAP-NLG in the same neuron (indicated by overlap of green and blue markers) resulted in diffuse cis ID-PRIME signal (row 1) much stronger than the trans-cellular ID-PRIME signal (row 2) in the same dish. Trans-cellular ID-PRIME signal was always localized to contact sites (arrow heads). Omission of lipoic acid suppressed both cis (row 3) and trans (row 4) ID-PRIME signal. Scale bars, 10 μm.

We also used sequential lipofection of plated hippocampal neurons to test picolyl azide ID-PRIME. Figure 6-15A shows fluorescent signal from the Alexa Fluor 647-alkyne (AF647-alkyne) used for picolyl azide detection at contact sites between blue LpIA-NRX-expressing neurons and green LAP-NLG-expressing neurons.
neurons. A negative control with picolyl azide omitted showed no labeling. The trans picolyl azide ID-PRIME signal was also localized to contact sites, but was weak – considerably weaker than lipoic acid ID-PRIME signal with the same reporter constructs, perhaps because the two-tiered antibody detection of lipoic acid offers signal amplification. Accordingly, we were unable to detect picolyl azide ID-PRIME signal in nucleofected neurons which express the reporters at lower levels than lipofected neurons (data not shown).

Since the fusion site of BirA in NRX3β influenced the efficiency of BLINC, we also prepared a stalk-domain fusion of *LplA in NRX3β at the same site as BirA, (*LplA_{272}-NRX3β) and tested this construct for picolyl azide ID-PRIME in lipofected neurons. Figure 6-15B shows that *LplA_{386}-NRX3β and *LplA_{272}-NRX3β gave comparable ID-PRIME signals, suggesting that ID-PRIME is less sensitive to the fusion geometry of *LplA, possibly because both fusion constructs were expressed more abundantly than any BirA-NRX3β construct in neurons.
Conclusions

In summary, our work presents three findings: (1) The BLINC methodology introduced in previous work (Thyagarajan and Ting, 2010) (paper now retracted) could be reproduced in HEK cells, but not in neuron cultures using the constructs and protocols previously described. (2) By re-designing the NRX and NLG fusion constructs and modifying the labeling protocol, we were able to
achieve BLINC labeling in neuron cultures. (3) Due to poor surface expression of BirA fusion constructs, our new BLINC protocol was still not robust in neurons, so we developed ID-PRIME for detection of trans-cellular NRX-NLG interactions. ID-PRIME was much more robust than BLINC in neuron cultures and the signal could be read out by either antibodies or small-molecule reagents.

Regarding the first point, the reasons for the lack of reproducibility of BLINC in neurons that we observed were: (i) the CMV promoter plasmids used previously (Thyagarajan and Ting, 2010), when introduced at DIV0 by nucleofection, expressed only transiently and could not be detected by DIV5. All previously reported experiments were performed in neurons at DIV5 and DIV16 (Thyagarajan and Ting, 2010), when BLINC reporters were no longer present. (ii) Even after changing the promoter from CMV to CAG to obtain persistent expression after nucleofection, the 1xAP tag on NLG1 was barely detectable above background. (iii) BirA inserted near the N-terminus of the mature NRX, as in the previous study, did not give detectable BLINC signal in neurons, even when the CAG promoter and a 3xAP-NLG1 were used. (iv) Use of biotin-AMP gave high background signal in neuron BLINC experiments. This reagent was used in all experiments in the previous study.

We were able to achieve BLINC labeling in neuron cultures by driving persistent expression with CAG promoters, installing 3xAP in place of 1xAP on NLG, moving BirA to the stalk domain of NRX, and using biotin+ATP instead of biotin-AMP. Nevertheless, we found that BLINC in neurons was not robust and sometimes failed, likely due to poor surface targeting of even our best BirA-NRX fusion. It is, however, a strength of the labeling system that such low, virtually non-detectable levels of surface BirA expression could produce detectable BLINC signal, attesting to the high sensitivity of streptavidin-fluorophore detection. Engineering of the BirA sequence or exploration of alternative fusion sites may improve surface BirA expression in future reporter designs.

In contrast to neurons, BLINC in non-neuronal cells (e.g., HEK) and in mixed neuron-HEK cultures, which comparatively exhibit higher surface expression of BirA constructs, was very reliable and specific. The signal was also interaction-
dependent, and well-localized to cell-cell contact sites, and should therefore be a useful method with which to detect and image other trans-cellular protein-protein interactions.

Lastly, we introduced in this study a new methodology for imaging trans-cellular protein complexes using interaction-dependent PRIME (ID-PRIME). In ID-PRIME, an LpIA mutant replaces BirA, and LAP replaces AP. The advantages of ID-PRIME over BLINC are two-fold: (1) The *LpIA-NRX fusion has much better surface targeting than BirA-NRX (despite its slightly larger size of 38 kD), leading to robust and reproducible signal in neuron cultures that is localized to NRX-NLG contact sites. (2) Using the picolyl azide detection strategy (Uttamapinant et al., 2012; White and Zegelbone, 2013a) (Figure 6-10), NRX-NLG interactions can be read out with small, bright fluorophores that are less likely to introduce trafficking artifacts or be sterically excluded from crowded synaptic regions compared to the streptavidin reagent used for BLINC. This work demonstrated ID-PRIME in three experimental configurations: HEK cultures (Figure 6-12), mixed HEK-neuron cultures (Figure 6-13B), and neuron cultures (Figure 6-13A). Given its advantages over BLINC, ID-PRIME expands the toolkit available for imaging trans-cellular protein-protein interactions in living cells.

There are still some issues to be resolved, however, before ID-PRIME can be a minimally-invasive and faithful tool with which to study physiologically relevant NRX-NLG interactions in neurons. The major concern is that its sensitivity must be improved to the point that signal can be easily detected even when reporter constructs are not overexpressed. This is particularly true for picolyl azide ID-PRIME, which currently has lower sensitivity than lipoic acid ID-PRIME in neurons even though its reagents are more suitable for reporting on synaptic protein complexes. A second concern is that we did not examine the effects of neurexin and neuroligin shedding on BLINC or ID-PRIME signal in this work. Recent studies have shown that the ectodomains of both NRX (Bot et al., 2011; Saura et al., 2011) and NLG (Peixoto et al., 2012; Suzuki et al., 2012) may be cleaved by membrane-anchored metalloproteases in a potentially activity-dependent manner. This could complicate the interpretation of BLINC and ID-
PRIME data if some fusion constructs of NRX and NLG are more prone to cleavage than wild-type (giving false negatives), or if labeling signal on the cleaved NLG ectodomain that ought to have escaped into the medium is trapped by full-length NRX (giving false positives). We plan to study and address these limitations in future work.

Materials and Methods

Cloning of plasmids

Constructs were prepared by standard restriction cloning methods and QuikChange mutagenesis (Stratagene). Overlap extension PCR was used to clone pCAG- BirA<sub>36</sub>-NRX3β and pCAG-BirA<sub>272</sub>-NRX3β.

HEK Cell Culture and Transfection

Human embryonic kidney 293T (HEK) cells were grown in minimal essential medium (MEM, Mediatech) supplemented with 10% (v/v) fetal bovine serum (PAA Laboratories) at 37°C under 5% CO<sub>2</sub>. Cells were typically transfected at ~70% confluence using Lipofectamine 2000 (Life Technologies) using the manufacturer’s suggested protocol. Cells for imaging were grown on 150 μm glass coverslips pre-coated with 50 μg/ml human fibronectin (Millipore). Approximately 24 hours after transfection, cells were lifted by trypsinization, co-plated at ~80% density, and labeled for BLINC or ID-PRIME ~24 hours later.

Rat Hippocampal Neuron Culture and Lipofection

Sprague Dawley rat embryos were sacrificed at embryonic day 18. Dissected hippocampal tissue was digested with papain (Worthington) and DNasel (Roche), then plated on 0.09–0.12 mm thickness glass coverslips (Carolina
Biological Supply) in a 1:1 volume ratio of growth medium A and growth medium B and cultured at 37°C under 5% CO₂. Growth medium A is MEM (Sigma) with L-glutamine (Sigma) supplemented with 10% (v/v) fetal bovine serum (PAA laboratories) and 2% (v/v) B27 (Life Technologies). Growth medium B is Neurobasal medium (Life Technologies) supplemented with 2% (v/v) B27 and 1% (v/v) GlutaMAX (Life Technologies). Glass coverslips were pretreated with poly-D-lysine (Sigma) and mouse laminin (Life Technologies). At 2 days in vitro, half of the spent culture medium was replaced with fresh growth medium B. This process was then repeated every 24 hours starting at 5 days in vitro. Transfection by Lipofectamine 2000 was performed between 5 to 11 days in vitro, using 1 μL Lipofectamine 2000 reagent per 1.91 cm² well (less than the manufacturer’s recommendation, to reduce toxicity).

**Neuron Nucleofection Using the Rat Neuron Nucleofector Kit (Lonza)**

Dissociated neurons were suspended in 100 μL nucleofection solution provided by the kit at a density of 1–1.5×10⁷ cells/mL, and mixed with 1 μg of each plasmid construct, then transferred into a nucleofection cuvette and nucleofected with the O-003 program. Cells were rescued in growth medium A (pre-warmed to 37°C), then plated onto poly-D-lysine treated glass coverlips (12 mm diameter) at a density of 150,000 cells per 1.9 cm².

**BLINC Labeling of HEK and Neuronal Cultures**

BLINC labeling was typically carried out 24 hours after co-plating for HEK cell cultures, or 5–12 days after co-plating for neuron cultures. Cells were incubated in growth medium B containing 20 μM biotin (gift from Tanabe USA), 500 μM ATP, and 1.25 mM magnesium acetate for 5–15 min. at 37°C. Cells were then rinsed three times with Tyrode’s buffer (145 mM NaCl, 1.25 mM CaCl₂, 3 mM KCl, 1.25 mM MgCl₂, 0.5 mM NaH₂PO₄, 10 mM glucose, 10 mM HEPES, pH 7.4) and subsequently stained with wild-type streptavidin-Alexa Fluor 647 (AF647) or
monovalent streptavidin-AF647 conjugate (Howarth et al., 2006a) in Tyrode’s buffer supplemented with 0.5% (w/v) vitamin-free casein (MP Biomedicals) for 5 min. at 37°C. Cells were rinsed three more times with Tyrode’s buffer before imaging. Alternatively, HEK cells were treated with 5 µM biotin-AMP (Howarth and Ting, 2008) for 2 min. at 37°C instead of biotin plus ATP (in Figure 6-9). We found that labeling with biotin-AMP was suitable for HEK cells but not for neurons, as it produced high background (in Figure 6-9). Biotinylation of total surface AP using purified enzyme was performed in the same way except that 1 µM biotin ligase (Howarth and Ting, 2008) was also added during the first labeling step.

**ID-PRIME Labeling of HEK and Neuronal Cultures**

ID-PRIME labeling was typically carried out 24 hours after co-plating for HEK cell cultures, 5–12 days after co-plating for nucleofected neuron cultures, or 1–2 days after sequentially lipofecting neurons. Cells were treated with Tyrode’s buffer containing 500 µM ATP, 1.25 mM magnesium acetate, and either 100 µM DL-α-lipoic acid (Alexis Biochemicals) or 100 µM picolyl azide (Uttamapinant et al., 2012) for 15–20 min. at 37°C.

To detect lipoic acid, cells were rinsed three times in Tyrode’s buffer and subsequently stained with a 1 : 200 dilution of rabbit anti-lipoic acid polyclonal antibody (Calbiochem) for 5 min. in the same buffer. Cells were again rinsed three times, followed by a 1 : 300 dilution of goat anti-rabbit secondary antibody AF647 conjugate (Life Technologies) for the same time in the same buffer. Cells were imaged live after three further rinses.

To detect picolyl azide on HEK cells, cells were rinsed three times in Tyrode’s buffer and treated with 50 µM CuSO₄, 2.5 mM sodium ascorbate, 20 µM alkyne-AF647 conjugate (Life Technologies), 250 µM THPTA ligand (Chan et al., 2004) and 100 µM 4-hydroxy-2,2,6,6-tetramethylpiperidin-1-oxyl (TEMPOL,
Calbiochem) in the same buffer for 5 min. at room temperature. Ligand bound 
Cu(I) complexes were pre-formed by mixing the CuSO₄, sodium ascorbate and 
THPTA, and incubating at room temperature for 10 min., before the alkyne and 
TEMPOL were added. Cells were rinsed three times further, then fixed with 4% 
(v/v) formaldehyde in a pH 7.0 buffer containing 0.12 M sucrose for 15 min. at 
room temperature before imaging.

Azide detection in neurons was performed after fixation. After picolyl azide 
ligation on living neurons as described above, neurons were fixed with 4% (v/v) 
formaldehyde in Tyrodes buffer, then blocked in Tyrode’s buffer supplemented 
with 0.5% (w/v) casein for one hour. Cells were treated with 1 mM CuSO₄, 2.5 
mM sodium ascorbate, 5 µM alkyne-AF647 conjugate, 100 µM TBTA ligand 
(Chan et al., 2004) and 100 µM TEMPOL in the same buffer for one hour at room 
temperature. Ligand bound Cu(I) complexes were pre-formed by mixing the 
CuSO₄, sodium ascorbate and TBTA, and incubating at room temperature for 10 
min., before the alkyne and TEMPOL were added. Cells were imaged after three 
进一步 rinses.

**Immunofluorescence Detection of BirA and LplA Fusion Constructs in HEK 
and Neuronal Cultures**

For live-cell immunofluorescence detection, cells were incubated with a 1 : 200 
dilution of either a mouse anti-c-Myc antibody (Calbiochem) or a rabbit anti-HA 
antibody (Rockland) in Tyrode’s buffer supplemented with 0.5% (w/v) casein for 
15 min. at 37°C. Cell were rinsed three times with Tyrode’s buffer, and 
subsequently stained with the corresponding secondary antibody: goat anti-
mouse-AF568 conjugate or goat anti-rabbit-AF568 conjugate (Life Technologies) 
in the same buffer. Cells were rinsed three times with Tyrode’s buffer before 
imaging live at room temperature.
For immunofluorescence detection on fixed cells, samples were fixed with 4% formaldehyde in Tyrode’s buffer, then permeabilized with methanol at −20°C. Cells were blocked for one hour in Tyrode’s buffer supplemented with 0.5% (w/v) casein, followed by primary antibody detection for one hour in the same buffer. A 1 : 200 dilution of one of the following antibodies was used: mouse anti-c-Myc antibody (Calbiochem), rabbit anti-HA antibody (Rockland), or mouse anti-FLAG M2 antibody (Agilent). Cells were then rinsed three times with Tyrode’s buffer and subsequently stained with the corresponding secondary antibody: goat anti-mouse-AF488 conjugate, goat anti-rabbit-AF488 conjugate, or goat anti-rabbit-AF568 conjugate (Life Technologies). Cells were rinsed three times further with Tyrode’s buffer before imaging.

Confocal Fluorescence Microscopy

Neuron cultures placed in Tyrode’s buffer or HEK cells placed in Dulbecco’s phosphate buffered saline (Gibco) were imaged using a Zeiss AxioObserver.Z1 inverted confocal microscope with a 40X oil-immersion objective. The microscope was equipped with a Yokogawa spinning disk confocal head, a Quadband notch dichroic mirror (405/488/568/647 nm), and 405 (diode), 491 (DPSS), 561 (DPSS), and 640 nm (diode) lasers (all 50 mW). BFP (405 nm laser excitation, 445/40 emission filter), GFP/Venus/AF488 (491 nm laser excitation, 528/38 emission filter), dsRed/tdTomato/AF568 (561 nm laser excitation, 617/73 emission filter), AF647 (640 nm laser excitation, 700/75 emission filter), and DIC images were collected using a Cascade II:512 camera and processed using SlideBook software version 5.0 (Intelligent Imaging Innovations). Acquisition time ranged from 10–2000 milliseconds. Neuron images in Figures S4 and S5 were projection summations from 0.5 μm-step optical stacks spanning 3.5 μm total depth.

Quantification of Lipoic Acid ID-PRIME Sensitivity in Neurons
Analysis was performed on 10 fields-of-view using the SlideBook software. For each field-of-view, one binary mask was created for each of the two fluorescent protein transfection markers, Venus and tdTomato. The two masks were then intersected to create an intersection mask, which ranged from 26 to 172 puncta or oblong segments (totaling 741 from 10 fields-of-view). Maximum lipoic acid ID-PRIME pixel intensity was tabulated for the 741 puncta or segments, and those with ID-PRIME signal-to-noise ratio greater than 3 : 1 were tallied, giving 38%. If the signal-to-noise ratio requirement was relaxed to 2 : 1, then 54% of Venus/tdTomato overlaps were positive for ID-PRIME signal. Noise was defined as the averaged ID-PRIME intensity on three non-transfected cells. Microscope instrument noise, defined as the ID-PRIME intensity on an area of the glass coverslip with no cell coverage, was subtracted from both signal and noise before the signal-to-noise ratio was calculated.

**Chemicals and Reagents**

ID-PRIME reagents for picolyl azide labeling are described in reference (Uttamapinant et al., 2012; White and Zegelbone, 2013b). All chemicals were purchased from Sigma-Aldrich unless otherwise specified.

**References**


Peixoto, R.T., Kunz, P.A., Kwon, H., Mabb, A.M., Sabatini, B.L., Philpot, B.D.


Chapter 7: Engineering lipoamidase as an eraser for LpIA dependent labeling
Introduction

The advent of green fluorescent protein GFP (Tsien, 1998) has paved the way for an explosion of biological studies using fluorescence imaging to study proteins in live cells. Although GFP has a distinct advantage over organic fluorophores of being genetically encodable and hence absolutely specific for labeling proteins inside cells, it suffers from a number of disadvantages, 1) it is relatively large (27 kDa) compared to organic dyes and this might interfere with the protein's natural function and trafficking, 2) restricted in its fusion sites, typically N- or C terminally fused, while organic dyes can be targeted to certain sites within the protein using the appropriate conjugation methods (Fernández-Suárez and Ting, 2008), although this has its caveats as well, and lastly 3) confers only one functionality (fluorescence) and 4) has low brightness compared to many organic fluorophores (Fernández-Suárez and Ting, 2008).

To obtain the synergistic advantages of both organic dyes while still being able to label proteins specifically, the Ting lab has engineered E.coli enzyme lipoic acid ligase (LpLA) (Fernández-Suárez et al., 2007; Liu et al., 2014; Uttamapinant et al., 2010; Yao et al., 2012) through site-directed mutagenesis to incorporate a variety of unnatural small molecules containing various functional handles/fluorophores site-specifically onto the lysine residue in a 13-amino-acid LpA Acceptor Peptide (LAP) tag (Puthenveetil et al., 2009) that is genetically fused to the protein of interest, similarly to GFP. Expression of the evolved Lpla enzyme within cells together with the protein of interest fused to LAP allows for intracellular labeling of the protein with various probes that can be added exogenously in the culturing medium.

Covalent labeling of proteins with small molecules is powerful because the modification adds a dimension of versatility towards manipulating the protein and the probes themselves are small and less perturbing to the natural function of the protein than fluorescent proteins. Functional handles like biotin can be used for pull-downs with streptavidin (Green, 1975) or fluorescent probes can be used to monitor a protein's trafficking inside cells. Covalent labeling however is
permanent, and the protein cannot be re-probed to obtain new information using a different modification after the initial labeling has occurred. To overcome this problem, we propose a way to reversibly label proteins covalently by engineering an enzyme that carries out the reverse reaction analogous to an "eraser".

LpIA, whose natural function in E. coli is the ligation of the cofactor lipoic acid onto one of three substrates (Cronan et al., 2005; Green et al., 1995), E2p subunit of pyruvate dehydrogenase (PDH), E3 subunit of 2-oxoglutarate dehydrogenase (OGDH), and H protein of glycine cleavage system, is able to catalyze both the lipoic acid adenylation and ligation step. Since the Ting lab had already co-opted and engineered LpIA towards the ligation of various functional probes onto the E2p and LAP substrates, we envisioned that by a similar analogy an enzyme that catalyzes the reverse reaction; removal of lipoic acid from its lipoyl domains, might be open towards engineering of its removal of the same fluorophores/functional handles LpIA has been engineered to ligate onto LAP (Figure 7-1).

![Figure 7-1. Example Scheme for reversible labeling to incorporate different functional handles. A blue fluorophore ligase first catalyzes the covalent ligation of a blue fluorophore probe onto the lysine residue of a 13-aa recognition peptide (LAP). To probe the same protein with a different functional handle, LAP is next erased by an enzyme that catalyzes the removal of the functional handle, followed by the ligation of a red fluorophore (or other functional handle depending on the user's needs) onto LAP.](image)

**Identifying an enzyme to reverse LpIA dependent labeling**

*E. faecalis* lipoamidase (Lpa) is an 80.6 kDa enzyme first observed by Reed and co-workers (SUZUKI and REED, 1963) that was able to cleave lipoic acid from a number of small lipoalyzed substrates and the lipoyl domain of pyruvate dehydrogenase complex. Although discovered more than 50 years ago, the gene encoding Lpa was only recently cloned and purified to homogeneity by...
Cronan and co-workers (Jiang and Cronan, 2005). This enzyme was revealed to belong to the Ser-cis Ser-Lys amidohydrolase (Labahn et al., 2002; Shin et al., 2003) family, and exists as a dimer in solution, similar to malonamidase E2 (Shin et al., 2002) and fatty acid amide hydrolase FAAH, both amidohydrolases. This family of amidases has a unique Ser-Ser-Lys catalytic triad (Fig. 7-2), different from the classical Ser-His-Asp triad commonly found in nature. The highly conserved and unusual cis conformation of the second Ser in all Ser-Ser-Lys amidases places both the amide nitrogen and its side chain at ideal positions to polarize the nucleophilic Ser (based on the FAAH (1MT5 in Protein Database Bank (PDB)) and E2 (PDB ID:1OCM) structures) and we can extrapolate that this is repeated within Lpa as well based on its highly homologous amidase domain.

Lpa is postulated to have a scavenging role for its host organism, which is a lipoate auxotroph lacking LipA and LipB (Jordan and Cronan, 2003; Miller et al., 2000; Morris et al., 1995), (the enzymes used in the biosynthesis of lipoate), but possesses two Lpla homologues that could be used to ligate the scavenged lipoic acid onto its three targets as described earlier.

Lpa was shown to have some activity for releasing lipoic acid bound to lysine, and also a much reduced but low activity for biotinyl-lysine (Jiang and Cronan, 2005). This promiscuity lead us to believe that Lpa’s active site might be malleable towards engineering, specifically towards removing probes that our lab had engineered the ligase towards using. We decided to begin by first isolating
the enzyme to characterize Lpa’s substrate specificity, and using a rational design evolution platform, screen for possible mutations that could improve and expand Lpa’s substrate scope.

In order to carry out reversible labeling, the “eraser” would ideally possess the following criteria, 1) remove the probe of interest efficiently within the time frame of the biological process studied (high \( k_{cat} \)), 2) removal of the probe must maintain the integrity of the protein of interest such that it can be relabeled, 3) removal of the probe must occur efficiently at physiologically relevant concentrations of the protein of interest studied (low \( k_m \)), and, 4) the protein should function optimally under the conditions determined within the cell’s milieu and have little effect on endogenous cellular activity. It is unknown if Lpa has any endogenous targets when expressed in mammalian cells, we hypothesize that because the predominant lipoyl containing proteins in mammalian cells are the E2 subunits of PDH and α-ketoglutarate dehydrogenase KGDH complexes (Berman et al., 1981; Stepp et al., 1981), which are localized to the mitochondrial matrix and unavailable in the cytosol, it is unlikely that Lpa would have any sort of deleterious effects on the mammalian cell. We investigated Lpa’s potential for fulfilling these criteria as well as how to improve its properties via engineering.

**Purification of Lipoamidase**

Attempts by Cronan and co-workers to express Lpa in BL21-DE3-pLysS failed despite extensive exploration of various growth conditions, including addition of acetate and succinate to bypass inactivated PDH and OGDH, owing to leaky expression of the protein (Jiang and Cronan, 2005) even when the cells were under tighter transcriptional regulation of the T7 promoter. Lpa is toxic in *E.coli* because it removes the lipoic acid from complexes used in oxidative decarboxylation, impairing the citric acid cycle (Jiang and Cronan, 2005). To express the protein in *E.coli TOP10* cells that lack the T7 polymerase gene, T7 RNA polymerase was exogenously provided via infection of the host strain carrying the Lpa expression plasmid with phage λCE6 at \( OD_{600} = 0.4 \).
I attempted this several times under different OD induction, MOI conditions, but could not see any improvement in the over-expression protein lysate pattern.

*Expression and attempts at purification of Lpa using λCE6 induction*

![Image](image_url)

**Figure. 7-3** λCE6 induction mediated induction of XL1B containing cells transformed with the Lpa containing plasmid. TOP10 cells were transformed with an expression plasmid for Lpa, and induced with λCE6 phage to introduce T7 polymerase into cells. Lanes are: uninduced (UI), induced (I), lysate supernatant (LS), flowthrough (FT), binding with 0mM imidazole (B), wash with 30mM imidazole (W), elution with 200mM imidazole (E).

**Co-expression of Lpla and Lpa in BL21 cells rescues Lpa lethal phenotype**

Because the lambda phage induction method, in both the Cronan lab and our hands (Fig. 7-3), yielded poor expression of Lpa and we asked if a different method of overcoming Lpa’s toxicity could be used to purify it. We hypothesized that co-expression of Lpla, supplemented with lipoic acid, could rescue the toxicity associated with expression of Lpa in *E.coli,* since LplA would re-ligate lipoic acid back into its lipoyl domain targets. We cloned an untagged LplA overexpression plasmid and co-transformed this with the His-tagged Lpa expression plasmid in BL21-DE3 cells, selecting for cells that expressed contained both plasmids (under two different antibiotic selection markers, ampicillin and kanamycin). Fortuitously, the lethal phenotype previously associated with cells transformed with the Lpa plasmid was not observed (no colonies are observed when this plasmid is transformed alone) and we proceeded to use this system to try to purify Lpa protein by Ni-NTA affinity purification (Fig.7-4).
Figure 7-4. Co-expression of Lpla rescues the Lpa lethal phenotype in BL21DE3 cells.

Overexpression and purification of Lpa coexpressed with Lpla in Bl21 cells

This co-expression system was then applied towards purifying Lpa on a larger scale, 1L Luria broth (LB) culture. This was supplemented by 10 ug/ml lipoic acid, 5mM sodium succinate, 10mM MgSO₄, 5mM sodium acetate and 0.2% maltose. Cells were induced with IPTG at OD₆₀₀ = 0.4 and allowed to continue at room temperature for 20 hours. The following steps occurred at 4 degrees Celsius. The cells were pelleted and lysed with BPER lysis reagent in the absence of protease inhibitor cocktail and PMSF, because these reagents would inactivate the hydrolase activity of Lpa (which is a hydrolase). The lysate was pelleted again and the supernatant applied to a Ni-NTA column and eluted over a gradient of 0-200mM imidazole elution buffer. A protein with a molecular weight of approximately 100kDa as verified by SDS PAGE was retained and eluted off the Ni-NTA column at 200mM imidazole elution buffer, and this was shown by Western blotting to have a His₆ tag and likely to be the Lpa protein. A number of degradation products containing His₆ tags were also seen and both this observation and the apparent higher molecular weight of Lpa (predicted MW: 80.6 kDa) correlated with the Cronan Lab's observation during their Lpa purification. Although we were unable to purify Lpa to a completely pure state in high yield, we proceeded to apply the protein towards testing its substrate specificity in vitro, in order to test Lpa's substrate promiscuity and see if it would be viable for use as an “eraser”.
Figure 7-5. Overexpression and purification of Lpa by co-expression with LpIA in BL21 cells. SDS-PAGE gel of Ni-NTA purification steps of Lpa coexpressed with Lpla. Cells were grown in Luria Broth (LB) supplemented with 5mM succinate and 5mM sodium acetate until \( \text{OD}_{600} \) reached 0.6, cells were then induced with IPTG for 20 hours at 30°C. Cells were lysed with BPER reagent (Thermo) and then incubated with Ni-NTA resin for purification. The last lane (marked 14) shows an anti-His tag blot of the last elution aliquot showing full length protein and degradation products (no protease inhibitors were used because these inhibit Lpa activity). Lpa is supposed to be an 80.6kDa protein by calculated molecular weight, but behaves like a ~97kDa protein as observed by SDS PAGE. This observation mimics that seen by Cronan and co-workers (Jiang and Cronan, 2005). Lanes are: uninduced (UI), induced (1), lysate supernatant (LS), flowthrough (FT), binding with 0mM imidazole (B), wash with increasing concentrations of imidazole, starting from 30mM to 200mM imidazole.

Testing Lpa activity on different substrates in vitro

We first applied Lpa towards removing lipoate from a lipoylated E2p substrate generated by a similar co-overexpression of E2p with LpIA in BL21 cells, which was also supplemented with lipoic acid to label the 9 kD lipoate acetyltransferase E2p subunit in the pyruvate dehydrogenase complex (Green et al., 1995). We were pleased to find that Lpa was able to remove lipoic acid from E2p as expected from its native function. Since we wanted to apply Lpa towards removing substrates ligated to the LpIA acceptor peptide LAP2, an evolved 13-amino acid peptide substrate for LpIA (Puthenveetil et al., 2009), we next tested if Lpa could remove different functional handles that had been pre-ligated onto the LAP2 peptide. We first tried to apply Lpa towards 7'OH coumarin ligated to the
LAP2 peptide. Since the lab had recently engineered an intracellular fluorophore ligase (Uttamapinant et al., 2010); the LpIA mutant W37V that could ligate 7’OH coumarin (structure (3) in Fig. 7-6) to proteins fused to the LAP2 peptide, we thought this would be an ideal target for engineering an “eraser” to. We were disappointed to find that Lpa was unable to remove coumarin from LAP2 and thought perhaps that engineering of the active site of Lpa, as had been applied towards engineering the active site of LpIA towards ligating coumarin, might allow Lpa to eventually remove coumarin from LAP2. Lpa was however able to remove both 7’azido-heptanoic acid (structure (1) in Fig. 7-6) and 10’bromo decanoic acid (structure (2) in Fig. 7-6) from LAP2, two biorthogonal probes that are functional substrates for Halotag ligation (Los et al., 2008), and copper catalyzed azide-alkyne “click”, Huisgen 1’3 dipolar cyloaddition (Chan et al., 2004) respectively. These functional handles have been applied previously towards fluorophore ligation in the Ting lab (Liu et al., 2012; Yao et al., 2012). These in vitro results suggest that Lpa is able to recognize long hydrophobic substrates that resemble lipoic acid, but larger substrates such as coumarin might not fit within its binding pocket.

![Chemical structures of functional handles that we would like to remove from prelabeled LAP2.](image)

Figure 7-6. Chemical structures of functional handles that we would like to remove from prelabeled LAP2. (1) 7’azido-heptanoic acid (2) 10’bromo-decanoic acid (3) 7’hydroxy-coumarin.
Figure 7-7. Lpa activity on lipoylated E2p. Lpa activity was tested in vitro on a lipoylated E2p substrate. 20nM of Lpa almost completely removes lipoic acid from E2p when added to lipoylated E2p (6.25uM) in dPBS over 1 hour. These were run on the HPLC with an acetonitrile gradient of 40%-57% over 20 minutes.

Figure 7-8. Lpa is able to remove lipoic acid from lipoylated LAP2 peptide. Lpa activity was tested in vitro on a pre-lipoylated LAP substrate. 20nM of Lpa almost removes lipoic acid from E2p when added to 15.65uM lipoylated LAP in dPBS over 1 hour. These were run on the HPLC with an acetonitrile gradient of 30%-60% over 10 minutes.
Figure 7-9. Lpa activity on coumarin ligated LAP2 peptide. Lpa activity was then tested against a coumarin ligated LAP2 peptide (made in vitro). No coumarin removal was observed after incubation for 1 hour, with 200nM Lpa and 15.65uM 7-hydroxy-Coumarin-LAP2 peptide. These were run on the HPLC with an acetonitrile gradient of 30%-60% over 10 minutes.
Figure 7-10. Lpa activity on Azide7 or 10Br-decanoic acid ligated LAP2 peptide. Lpa activity was also tested against the Azide7 or 10-Bromo-decanoic acid ligated to the LAP2 peptide (made in vitro). Azide7 and 10-Bromo-decanoic acid removal was observed after incubation for 1 hour, although this was incomplete, with 200nM Lpa and 15.65uM Azide7/10-Bromo-decanoic LAP2 peptide. These were run on the HPLC with an acetonitrile gradient of 30%-60% over 10 minutes.

**Testing Lpa activity on the cell surface of HEK cells**

Since we would eventually apply Lpa towards labeling proteins in cells, we tested the activity of exogenously added Lpa on HEK cells expressing E2p on the cell surface. HEK 293T cells were transfected with E2p-CFP-TM (E2p fused to cyan fluorescent protein and a transmembrane domain from PDGFR) and mCherry-Lpla-ER (Lpla fused to a KDEL sequence which targets it to the secretory pathway). These cells were incubated with lipoic acid supplemented within the culture media overnight. E2p expressed within the secretory pathway is labeled intracellularly with lipoic acid by the endoplasmic reticulum targeted Lpla. In the absence of Lpa, an anti-lipoic acid antibody detects lipoylated surface expressed E2p. Incubation of cells with 1.65uM Lpa for 1 hour followed by 3 washes in dPBS showed complete removal of lipoic acid, again by anti-lipoic acid staining (Fig. 7-11). We found that removal of lipoic acid from E2p substrate could be shortened to a 5 min incubation and 200nM Lpa for complete removal.
Figure 7-11. Lpa removal of lipoic acid from lipoylated-E2p expressed on the cell surface of HEK cells. E2p-YFP on the surface of live HEK cells was ligated with lipoic acid using 10 μM WT LpIA and 1 mM ATP for 15 min. Following, the cell media was replaced with 200 nM Lpa or dPBS to remove lipoic acid for 1 hour at room temperature. Cells were washed, then stained for lipoic acid live using anti-lipoic acid antibody. Negative controls are shown with lipoic acid omitted, DIC = differential interference contrast image.

Figure 7-12. Lpa removea lipoic acid from LAP expressed on the cell surface of HEK cells. LAP2-LDLR-GFP on the surface of live HEK cells was ligated with lipoic acid using 10 μM WT LpIA and 1 mM ATP for 15 min. Following, the cell media was replaced with 200 nM Lpa or dPBS to remove lipoic acid for 15 min at room temperature. Cells were washed, then stained for lipoic acid.
acid live using anti-lipoic acid antibody. Negative controls are shown with lipoic acid omitted during the LpIA step, DIC = differential interference contrast image.

We next tested Lpa’s activity towards removing lipoic acid when the LAP2 peptide was the lipoylated substrate. Lpa removed lipoic acid from LAP2-LDLR-GFP expressed on the cell surface, which had been labeled a priori with lipoic acid, followed by removal initiated by addition of 200nM Lpa and incubation for 15 minutes. This verified that Lpa not only recognized the endogenous E2p substrate, but also the evolved minimal variant LAP2 peptide (Fig. 7-12).

![Coumarin GFP DIC](image)

Figure 7-13. Lpa cannot remove 7'OH coumarin from LAP expressed on the cell surface of HEK cells. LAP2-LDLR-GFP on the surface of live HEK cells was ligated with 7'OH coumarin using 10 uM W371 LpIA and 1 mM ATP for 15 min. Following 3 washes in dPBS, the cell media was replaced with 200 nM Lpa or dPBS for 15 min at room temperature. Cells were washed and imaged live. Negative controls are shown with LpIA omitted during the labeling step.

We next focused on a differential functional moiety labeled on the same substrate, the LAP2 peptide fused to the N terminus of the low density lipoprotein receptor fused to GFP; LAP2-LDLR-GFP. LAP2 was now labeled with a more challenging functional group, the larger and less hydrophobic (as compared to the long alkyl chain substrates) 7-hydroxy-coumarin using exogenously added W371 LpIA. Addition of 200nM Lpa and incubation for 15 minutes did not result in a decrease in the coumarin channel’s signal (Fig. 7-13), verifying the in vitro data that showed that Lpa was not able to remove coumarin.
Figure 7-14. The LAP peptide can be reused for ligation after Lpa erasing of functional handles. LAP2-LDLR-GFP on the surface of live HEK cells was ligated with lipoic acid using 10 uM WT LpIA and 1 mM ATP for 15 min. Following, the cell media was replaced with 200 nM Lpa or dPBS to remove lipoic acid for 15 min at room temperature. Cells were washed, and either relabeled with lipoic acid using LpIA for 15 min again or just dPBS. Cells were then stained for lipoic acid live using anti-lipoic acid antibody (red channel). Negative controls are shown with lipoic acid omitted during the LpIA step, DIC = differential interference contrast image.

To answer if the LAP2 peptide which had undergone Lpa dependent lipoic acid removal was still amenable for relabeling with LpIA, we carried out an experiment in cells expressing LAP2-LDLR-GFP (using similar conditions for labeling and removal of lipoic acid as above), first labeling with lipoic acid by exogenously added LpIA, followed by 3 washes with dPBS and then erasing with Lpa, followed by 3 washes again and relabeling with lipoic acid and exogenous LpIA again. Lipoic acid signal was only observed in steps following exogenous LpIA and lipoic acid addition, but absent when Lpa was the last enzyme added to cells. This demonstrates that the LAP2 peptide remains “intact” after removal of its probe and can be re-probed by subsequent addition of a different functional handle (Fig. 7-14).
Lpa can remove both 10-Bromo-decanoic acid or 7'Azido-heptanoic acid from E2p

We next tested in cells Lpa’s wide substrate promiscuity as demonstrated in vitro (being able to remove a variety of alkyl chain substrates in addition to the endogenous target lipoic acid). Cells expressing E2p-CFP-TM were labeled with either 10-Bromo-decanoic acid, which is the substrate for Halotag labeling, or 7’Azido-heptanoic acid that can be derivatized by click chemistry with Cy5 alkyne conjugates respectively. 10-Bromo-decanoic acid was detected by Halotag conjugates to AlexaFluor568 (Liu et al., 2012), while 7’Azido-heptanoic acid was detected by copper catalyzed click with a Cy5-alkyne conjugate (Yao et al., 2012). Signal indicating that each probe was ligated to E2p, was only detected in the absence of the addition of Lpa, and we conclude that Lpa can remove those probes from E2p expressed on the surface of cells (Fig. 7-15 and 7-16). Lpa
however, was not able to remove 10-Bromo-decanoic acid when it was prelabeled with Halotag-AF568 conjugates (Fig. 7-15).

**Figure 7-16. Lpa can remove 7’azido-decanoic acid from E2p.** Cells transfected with E2p-CFP-TM were first labeled with Azide 7 using 10uM WT LpIA for 20 minutes then rinsed three times with dPBS. Cells in "erasing" lanes were then incubated for 5 minutes with 200nM Lpa, washed three times in dPBS. All cells were then labeled with Cy-5 alkyne conjugates for 10 min in the presence of 100uM CuSO₄ with 500uM THPTA as the ligand, and 2.5mM sodium ascorbate, then rinsed three times with dPBS. Cells were then imaged by Epi-fluorescence microscopy. Complete removal of Azide 7 by Lpa was observed.

**Figure 7-17. Lpa did not remove biotin from the biotin ligase acceptor peptide AP.** Cells expressing AP-CFP-TM were cotransfected with biotin-ligase BirA targeted to the secretory pathway (BirA-ER), and incubated overnight with 10uM biotin added to the culture media.
Labeling of AP with biotin by BirA occurs in the secretory pathway. 6uM WT-Lpa was added to cells for 15 min, and biotin was detected by streptavidin-647 conjugates. No detectable removal of biotin from the AP construct within the 15 minute time frame was observed.

Since Lpa was reported to remove biotinyl lysine (Jiang and Cronan, 2005), we tested Lpa’s ability to remove biotin from a biotinylated acceptor peptide (AP) (Fig. 7-17). HEK cells were transfected with biotin ligase fused to a KDEL sequence, BirA-ER, and the acceptor peptide fused to a cyan fluorescent protein and a transmembrane domain from PDGFR, AP-CFP-TM. Cells were cultured in media containing serum and an additional 10uM biotin overnight. The next day, cells were rinsed 4 times with dPBS to remove any free biotin and 6uM WT-Lpa was added for 15 minutes, before washing again three times, and detecting any biotinylated AP with WT-streptavidin-AF647 conjugate for 5 minutes. Under these conditions, we could not detect any differences between cells treated with or without Lpa.

**Engineering a coumarin amidase from Lpa**

We turned towards rational design to try to expand Lpa’s substrate specificity for 7'hydroxycoumarin. Although no crystal structures of Lpa exist, Lpa’s closest amidase relative by sequence alignment, is the well studied fatty acid amide hydrolase (FAAH) (McKinney and Cravatt, 2005), of which numerous crystal structures with various inhibitors bound were known (PDB ID: 1MT5, 4J5P, 4HB5). Lpa and FAAH can be aligned with a gap of only a single amino acid over residues 150-275 (Lpa’s numbering), the region containing the active site residues needed for hydrolysis of the amide linkage. Since most of FAAH’s targets are long chained aliphatic substrates such as anadamide (Deutsch and Chin, 1993), and many of the side chains lining the tunnel containing FAAH’s acyl substrates are hydrophobic both in FAAH and the homologous positions in Lpa, it seemed plausible that the two proteins might bind their substrates in a similar manner. Upon closer inspection of FAAH’s inhibitor bound crystal structure, we hypothesized that mutating large hydrophobic residues from the
FAAH structure, into alanines in the corresponding positions for Lpa, might expand the acyl substrate tunnel while maintaining the rigidity of the peptide backbone. This expansion could then be extrapolated into an enlargement of Lpa’s specificity, and possibly allow a larger substrate such as hydrocoumarin to be removed. The following table depicts the residues that we selected for alanine mutant screening (Fig. 7-18).

<table>
<thead>
<tr>
<th>FAAH</th>
<th>Lpa</th>
</tr>
</thead>
<tbody>
<tr>
<td>L192</td>
<td>W210*</td>
</tr>
<tr>
<td>F194</td>
<td>N212*</td>
</tr>
<tr>
<td>L380</td>
<td>L411*</td>
</tr>
<tr>
<td>F381</td>
<td>F412*</td>
</tr>
<tr>
<td>L404</td>
<td>I430</td>
</tr>
<tr>
<td>F432</td>
<td>S458</td>
</tr>
<tr>
<td>T488</td>
<td>T498</td>
</tr>
<tr>
<td>V495</td>
<td>Q502</td>
</tr>
</tbody>
</table>

Figure 7-18. Residues in FAAH that are less than 6 angstroms away from the MAP inhibitor. The corresponding residues were aligned and identified in Lpa and selected for alanine screening. 6 Angstrom zone around the MAP inhibitor (green) in FAAH (PDB ID: 1MT5). Large hydrophobic residues are colored in red (F), light blue (L), purple (V), yellow (W) and the phosphonate moiety of the inhibitor in orange.

**Alanine mutant screens for coumarin amidase activity**

These corresponding residues in Lpa were mutated to alanine by Quickchange mutagenesis, then expressed and purified by co-expression with LpIA (similar to normal expression and purification of WT-Lpa in E.coli). These mutants were then tested for amidase activity on the LAP2 peptide preligated with hydroxycoumarin. We screened the mutants by HPLC assays with two substrates, 7’hydroxycoumarin ligated LAP2 peptide for coumarin amidase activity, and a lipoylated E2p protein substrate for lipoamidase activity. Activity
assays consisted of reacting the Lpa mutant with the 7'-hydroxycoumarin ligated LAP2 peptide overnight, then run out on the HPLC (Fig. 7-19). Several mutants showed activity for removing hydroxycoumarin from LAP2 (Fig. 7-19). Although these assays were carried out overnight at room temperature, complete removal of coumarin from the LAP2 peptide was never observed, and the best mutants only removed approximately 50% of the coumarin. This might perhaps be explained by a loss of activity caused by denaturation of the protein at room temperature over long periods or residual proteases because of the lack of protease inhibitors used during purification. Residues that showed promising coumarin amidase activity were W210A, N212A and F412A. These residues were subsequently mutated into amino acids that either improved flexibility (glycine) or were small (but larger than alanine) and contributed to hydrogen bonding (serine) or hydrophobic (valine). These were then subjected to similar rounds of screening by HPLC on the two substrates, 7'-OH coumarin ligated LAP2 peptide and also the lipoylated E2p protein to screen for coumarin amidase and lipoamidase activity respectively. Additionally double and triple mutants of the residues with coumarin amidase activity by alanine screens were also generated to screen for synergistic effects. We found that the best mutant at removing hydroxycoumarin from LAP2 was the N212G mutant Fig. 7-20), which also removed hydroxycoumarin from coumarin pre-ligated to E2p (Fig. 7-21).
Figure 7-19. Lpa mutants can remove 7'OH coumarin from LAP in vitro. Lpa activity was tested against a coumarin-ligated LAP2 peptide (made *in vitro* see Methods for more details). Coumarin removal was observed after incubation overnight, with 200nM Lpa or mutants and 15.65μM 7'OH coumarin-LAP2 peptide in dPBS. These were run on the HPLC with an acetonitrile gradient of 30%-60% over 10 minutes.
Fig. 7-20 HPLC screening of focused coumarin amidase mutants
Mutant Lpa activity was tested against a coumarin-ligated LAP2 peptide. Coumarin removal was observed after incubation overnight, with 200nM Lpa mutants and 15.65uM 7'OH coumarin-LAP2 peptide. These were run on the HPLC with an acetonitrile gradient of 30%-60% over 10 minutes.

Figure 7-21. N212G Lpa mutant can also remove 7'OH coumarin from E2p in vitro.
Lpa activity was tested in vitro on a coumarin ligated E2p substrate. 200 nM of Lpa mutant removes 7'OH coumarin from E2p when added to 7'OH coumarin ligated E2p (6.25uM) in dPBS overnight. These were run on the HPLC with an acetonitrile gradient of 40%-57% over 20 minutes.

Based on these assays the corresponding residues from Lpa that showed activity for hydroxycoumarin in FAAH are shown in Fig. 7-22.

Figure 7-22. Corresponding amino acids of Lpa mutants, in FAAH substrate binding pocket (1MT5) that have coumarin amidase activity. The hydrophobic cavity contains a number of large hydrophobic residues surrounding the methylicholylfluorophosphonate inhibitor (MAP) bound to S241. Residues shown are homologous residues in Lpa that have been
show by alanine screening, and subsequently mutated to improve activity on 7'OH-coumarin ligated LAP/E2p.

Survival in the absence of coexpression with LpIA is an indicator of WT Lpa activity of alanine mutants

An interesting characteristic of all mutants that are active for the hydroxycoumarin probe maintained WT Lpa activity and had a lethal phenotype when transformed in BL21 cells without coexpression of Lpla just like the WT Lpa, suggesting that wild type hydrolase activity is necessary for maintaining its ability to remove the unnatural probe and that substrate specificity is not completely orthogonal to the wild type (Fig. 7-23 shows an example of this).

Figure 7-23. Screen for Lpa activity by co-expression with LpIA in BL21 cells. Lpa activity can be assayed by lethal phenotype observed when expression plasmid for Lpa is transformed into BL21-DE3 cells. Mutants that had lipoamidase activity induced a lethal phenotype when transformed, while non-active mutants did not. Coexpression of LpIA, shown in plates on the left always rescued this lethal phenotype.
Hydroxycoumarin can be removed from LAP2/E2p in cultured cells by the N212G Lpa mutant

The HPLC assays with 7’OH coumarin ligated LAP2 peptide/E2p revealed that the removal for hydroxycoumarin was best mediated by mutants of the N212 position in Lpa, although these assays also revealed that removal of hydroxycoumarin was much faster from E2p (removal was complete compared to LAP), suggesting that Lpa might recognize portions of the lipoyl domain of E2p that are not represented by LAP, which could contribute towards a higher amidase activity. The most active coumarin amidase mutants found were the N212G and N212A + W210A mutant. – approximately 50% removal overnight. Since most of the screening was carried out under overnight conditions, we tested the N212G mutant in a cell surface removal assay for coumarin amidase activity to see if it would be sufficiently active for experiments under physiological labeling and erasing conditions. HEK cells were first transfected with the following constructs, H2B-YFP (a transfection marker) and the E2p-CFP-TM (DEAD), a mutant that does not have CFP fluorescence since coumarin emits within the same wavelengths as coumarin. These were first labeled with 7’OH-coumarin by 10uM W37V LpIA for 30 minutes, and then rinsed thrice with culture media. Erasing with 2uM N212G Lpa was carried out for another 30 minutes then rinsed thrice with culture media. These were then imaged by fluorescence microscopy. All labeling and erasing occurred at 30 degrees Celsius and in DMEM+FBS. We found that hydroxycoumarin was only removed when the N212G mutant was used (Fig. 7-24). WT Lpa was used as a negative control since it did not have any in vitro activity for removing coumarin from the LAP2 peptide and was unable to eliminate coumarin signal.
Figure 7-24. Cell surface removal of hydroxycoumarin by the mutated lipoamidase. HEK cells expressing E2p on the cell surface (and H2B-YFP as a cotransfection marker) were first labeled using 10uM W37V Lpla coumarin ligase mutant for 30 minutes with 50uM hydroxycoumarin (3) and 1mM ATP. After three washes in dPBS, 2uM WT Lpa, the N212G mutant or just dPBS were added for 30 minutes, then washed with dPBS three times before imaging. Images show three fields of view, together with overlay of coumarin and YFP channels.

**Lipoamidase has resorufin amidase activity**

*Lpa activity towards resorufin (4) ligated to E2p*

Dr. Daniel Liu in the Ting lab, in collaboration with the Baker lab, had also developed a mutant LplA ligase (E20A, H147A and H149G) for a red fluorophore, resorufin (Liu et al., 2014), chemical structure is shown in Fig. 7-25. In tandem with developing a coumarin amidase, we thought we would also try developing a resorufin amidase using the mutants that we had screened to expand the active site pocket to tolerate larger substrates. An HPLC based screen of the removal of resorufin (Fig. 7-26) from preligated resorufin-E2p protein using various Lpa mutants revealed surprisingly that the wild type Lpa was able to remove the resorufin probe from the lipoyl domain E2p. This was surprising because of the
relatively larger difference in size of resorufin compared to the known substrates of Lpa, such as lipoic acid and other long alkyl acids. The wild type’s activity for resorufin despite having no activity for 7’OH coumarin was extremely surprising, and we tried to rationalize this observation. First resorufin does not have the same amide linkage in its alkyl linker that links the acid moiety to the fluorophore that is found in the 7’OH coumarin probe. This could have two effects, a hydrogen bonding surface that the wild type’s hydrophobic cavity might filter against, or a rigidity that might cause the large fluorophore moiety to clash with the hydrophobic tunnel. Additionally, Lpa was still a poor enzyme when the probe was ligated onto LAP, suggesting that Lpa did not recognize the LAP peptide as well as E2p.

We also tested resorufin amidase activity by cell surface removal assays. Resorufin was first ligated onto E2p-CFP-TM (dead) using the resorufin ligase (E20A, H147A and H149G) for 10 minutes at room temperature. Removal was Lpa dependent (Fig. 7-27) and could be inhibited by the addition of PMSF, which is a hydrolase inhibitor and inhibits Lpa activity.

![Figure 7-25. Chemical structure of resorufin.](image)
Figure 7-26. HPLC assays using various Lpa mutants with preligated resorufin-E2p protein with overnight incubation at room temperature. All mutants with WT Lpa activity were shown to be able to remove resorufin (as exemplified by Q502A mutant, which has impaired Lpa activity. Resorufin removal was observed after incubation overnight at room temperature, with 200nM Lpa and 6.25uM resorufin-E2p. These were run on the HPLC with an acetonitrile gradient of 40%-57% over 20 minutes. The +ATP+dPBS is a positive labeling control during the resorufin-E2p ligation step which no Lpa is added to, while -ATP+dPBS is a negative labeling control.
Figure 7-27. Resorufin labeling and removal by wild type Lpa. HEK cells expressing surface E2p were first labeled with resorufin (4) for 10 minutes using the resorufin ligase LpIA mutant (E20A, H147A and H149G) and 20uM resorufin probe (4). Cells were then washed 3 times in Tyrode's buffer. Cells were either "erased" with 3uM Lpa in either Tyrodes or in the presence of 1mM PMSF containing Tyrodes (an inhibitor of Lpa activity) for 15 min. Relabeling using the same conditions described earlier were carried out in one of the "erased" wells. Cells were then fixed in paraformaldehyde before imaging. All cells were fixed after the same amount of time to normalize for effects from endocytosis.

Conclusions

We now have a pair of fluorophore erasers that had semi-orthogonal activity, the N212G Lpa mutant for both the blue and red fluorophores, while the wild type worked solely for the red fluorophore. Both Lpa and its mutants are also able to remove functional handles such as the 7'azido-heptanoic acid and 10'bromo decanoic acid from E2p and LAP2. For Lpa to be useful, it is necessary to fulfill some of the criteria previously described. To measure the $K_{cat}$ and $K_m$ of the Lpa enzyme (criterion 1 and 3), it is necessary that we have a clean and homogenous source of protein. Since Lpa is difficult to express and purify sufficient quantities in E.coli as a result of its toxicity, we are looking at potential routes towards...
obtaining Lpa. We rationalize that because Lpa’s targets are localized within the mitochondria of eukaryotes, purification by cytoplasmic expression in yeast should not hinder growth or induce toxic effects and we might be able to improve our purification of the protein. Additionally, since we have tested Lpa’s application to cell surface proteins, it would be helpful to expand Lpa’s towards within cells. Intracellular labeling of a LAP-fluorescent protein fusion with lipoic acid by co-expressing LpIA could be used to assay for Lpa activity. Cells also expressing Lpa should have a much lower lipoic acid labeled signal if Lpa activity is present. Lpa’s large size might inhibit its potential to be used as a protein fusion that could be targeted to specific localizations. Truncating the protein sizably might allow for improved expression. Lastly, because Lpa is more active for the E2p substrate than LAP, evolving it towards recognizing LAP better is a very important step towards improving its usefulness. We hypothesize that a yeast surface display platform, cis-expressing Lpa and LAP could be used to evolve Lpa variants that recognize LAP, first by exogenous labeling of LAP with the probe of choice, followed by selecting for mutants that are best able to remove the probe most quickly.

Materials and Methods

Coexpression of LpIA and Lpa in BL21-DE3 cells
Chemical competent BL21-DE3 cells were made with the pCOLA-LpIA plasmid which is kanamycin resistant and LpIA is not fused to a HisX6 tag. These were transformed with the pYFJ202-His6-Lpa plasmid (ampicillin resistance) and plated on LB-plates containing both kanamycin and ampicillin to select for coexpression. BL21-DE3 cells transformed with just the Lpa plasmid showed a lethal phenotype while BL21-DE3 cells that had the LpIA construct formed many colonies, demonstrating that coexpression of LpIA could rescue the lethal phenotype.
HPLC screening to demonstrate Lpa removal of LpIA ligated probes on E2p or LAP2

E2p or LAP2 was first ligated (see below) with the probe of interest and quenched with 40mM EDTA (ethylenediaminetetraacetic acid, final concentration). For screens, reactions containing 200 nM Lpa enzyme, 15.65uM probe ligated LAP2 (sequence: H2N-GFEIDKVWYDLDACO2H), or 6.25uM probe ligated E2p in phosphate-buffered saline (PBS). Reactions were quenched with 1mM PMSF, and run on an HPLC with a C18 reverse phase column, recording absorbance at 210 nm. Elution conditions were 40%-57% over 20 minutes (E2p) or 30-60% acetonitrile in water with 0.1% trifluoroacetic acid over 10 min (LAP2) (linear gradient) at 1.0 mL/min flow rate.

In vitro probe ligation onto E2p or LAP peptide

In vitro probe ligation reactions contained 30uM LpIA enzyme (or mutant), 1mM probe, 80 uM E2p or 600uM LAP, 2 mM ATP, and 5 mM magnesium acetate in PBS and 10% glycerol were incubated at 30 °C for 4 hours before being quenched with 40mM EDTA (ethylenediaminetetraacetic acid, final concentration).

General protocol for cell surface probe labeling with LpIA and removal with Lpa

HEK cells plated on glass coverslips in wells of a 48-well cell culture plate (0.95 cm per well) were transfected with 100 ng pcDNA4-LAP2-LDL receptor or pNICE-E2P-CFP-TM using Lipofectamine 2000. At 18 hr-24hr after transfection, enzymatic ligation of the probe of interest (lipoic acid, 7'OH coumarin, resorufin, 7'Azido-heptanoic acid or 10-Bromo-decanoic) was performed in MEM with 10 uM LpIA (or mutant ligase), 200 uM probe, 1 mM ATP and 2 mM magnesium acetate for 15 min at room temperature (to minimize internalization of cell-surface proteins). After washing three times with MEM, cells were incubated with 200 nM Lpa for 15 min to remove ligated probe. Cells were then washed three times with MEM and imaged live. Resorufin and 7'OH coumarin were kindly provided by Dr. Daniel Liu or Dr. Chayasith Uttamapinant.
Detection of live cell azide ligation using cy5-alkyne via CuAAC.
After labeling and erasing steps described above, 7’Azido-heptanoic acid was detected by CuAAC by incubating cells with 50 μM Cy-5 alkyne conjugates for 10 min in the presence of 100μM CuSO₄ with 500μM THPTA as the ligand, and 2.5mM sodium ascorbate, then rinsed three times with dPBS before imaging. Cy-5 alkyne was kindly provided by Dr. Jennifer Yao.

Detection of live cell 10-Bromo-decanoic acid by HaloTag labeling
After labeling and erasing steps described above, 100nM of HaloTag-AlexaFluor 568 conjugate was added to cells for 5 min at room temperature, then washed three times with dPBS before imaging. HaloTag conjugated to AlexaFluor568 was kindly provided by Dr. Daniel Liu.

Protein purification of E2p
E2p in pET21a vector was transformed into E. coli BL21-DE3 strain. A single colony was picked from the ampicillin resistant plate and cultured in Luria Bertani (LB) medium supplemented with 100 μg/ml ampicillin at 37 °C and induced for expression with 420 μM isopropyl P-D-1-thiogalactopyranoside (IPTG) when OD₆₀₀ reached ~0.5. Cells were then transferred to a 21°C shaker and cultured for an additional 18 h. After pelleting, cells were lysed with B-PER reagent (Pierce) supplemented with protease inhibitor cocktail (Calbiochem) and 100 μM phenylmethylsulfonyl fluoride according to instructions from Pierce. Lysates were cleared by centrifugation at 8000 g for 15 min at 4°C. The supernatant was loaded onto a column containing Ni-NTA agarose resin (QIAGEN) and purified as for LpIA. Purified E2p was dialyzed against phosphate buffered saline (PBS), pH 7.4 and stored at -80°C.

Protein purification and the LAP2 peptide
LpIA and mutants were expressed and purified as previously described (Liu et al., 2014; Uttamapinant et al., 2010; Yao et al., 2012). The 13-amino acid LAP
peptide (H$_2$N-GFEIDKVWYDLDA-CO$_2$H) was synthesized by the Tufts University Peptide Synthesis Core Facility and purified to >96% homogeneity.

References


Green, N.M. (1975). Advances in Protein Chemistry Volume 29 (Elsevier).


